


```

2191 AAAAATGTAACCACTTCCAAACCGCGGATGAAGCTATCAAGAGT 2240
246 .LeuSerIlePheGluSerGluSplysGlyAspLysTyrAsnLysGlyT 262
2241 TCTTGGGCTTCCCAATACACATCAAGACAGCTTCAAT.....T 2284
262 yGlnAsnLeuLeuSerGlyLysValProThrLysProProthPro 278
2285 ATCAACAACACCCGTGAATTAGCA...GTTCCTGTTACACTCAAGAAAA 2331
279 GlyAspProMetProProAsnGlnProGlnThrThrserval.... 293
2332 ATCGACGATACCTTGGAAAAACAGCAACGCTCAAGCCATTAGATT 2381
294 ....LeuIleArgLysTyrAlaIleGlyAspTyrSerLysLeuGlu 309
2382 AAAAGTATCAAAAAATCTTCTCAGGTGAG.....AAAACCTTAGTG 2425
309 LysAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheGlnAlaArg 325
2426 GACCGCTTGTGATGTAGTGTAATAATGT.....CAACACACA 2466
326 ValPheSerSerAsnAsp.....IleGlyGluArgIleGlu 338
2467 TTAGTGGACAATAAGATGTACTATTCTTCCCAAAAGATGTGCGCT 2516
338 userAspGly...ThrTyrThrLeuThrGluLeuAsnSerProAlaGlyT 354
2517 ACAAAGGGGAAACGCTATACATTACTGAAAGTAAGCACTGCAGGAC 2566
354 ySerIleAlaGluProIleThrPheLysValGluAla.....Gly 367
2567 ATGAGTTAGCAAGAAACGACTGGCAAAVTGAGTGAGTGAGCAAGGC 2616
368 LysValLysThrIleLeuAspGlyLysGlnIleGluAsnProAsnLysG 384
2617 AAAGTA.....AGCATCGATGACAAAGAGTGCACCAATATCAAGT 2660
384 uIleValGluProTyrSerValGluAlaTyrAsnAspPheGluGluPhe 401
2661 TATT.....CCATTGGAAATTCGA.....AATAATTTCTTCTTTC 2688
401 eVal.....LeuThrThrGln.....AsnTyr 408
2699 CAATCAGAAATTAGAAATACACCATGCACAAATGGCAACAAGTAACT 2748
409 AlaLysPheTyrTyrAla.....LysAsnLysAsnLysSerG 422
2749 GCAGAGCGCAGCTTTTGGCTTGCACAAAGAAAAATGCTCAAGAACT 2798
422 nValValTyrCysPheAsnAlaAspLeuLysSerProProAspSerL 439
2799 AACTGTGGCAACTCAAAAAACAGAT..... 2823
439 spcLysGlyLysThrMetThrProAspPheThrGlyGluValLysTyr 455
2824 .....ACTACAGA..... 2832
456 ThrHisIleAlaGlyArgAspLeuPheLysTyrThrValLysProAla 472
2832 ..... 2832
472 pThrAspProAspThrPheLeuLysHisIleLysLysValIleGluLys 489
2833 .....TTGAGCTATTTT...AAAATAGTGAACCTG 2860
489 Lys...TyrArgGluLysGlyGlnAlaIleGluLysTyrSerGlyLeuThr 504
2861 GTGAGTATGCA.....ATGCTGACAAATCAGCA..... 2889
505 ThrGlnLeuArgAlaAlaThrGlnLeuAlaIleTyrTyrPheThrAsp 521
2890 ....CCATTAGCTACGACACTTGTGTAATATGAAATTACT..... 2931

```

```

521 rAlaGluLeuAspLysAspLysLeuLysAspTyrHisGlyPheGlyAsp 538
2932 .....GTTGATAA.....TATGGAAA. 2949
538 eIAsnAspSerThrLeuAlaValAlaLysIleLeuValGluTyrAlaGln 554
2950 .....ATTCACTATGCGAGGC 2964
555 .....AspSerAsnProGlnLeuThrAspLeuAspPhePheI 568
2965 AAAAATATTGAAGAAATGCGCGAATGACA..... 2997
568 eProAsnAsnAsnLysTyrGlnSerLeuIleGlyThrGlnThrHisPro 585
2997 ..... 2997
585 LysAspLeuValAspIleIleArgMetGluSplysGluValIlePro 601
2997 ..... 2997
602 ValThrHisAsnLeuThrLeuArgLys.....ThrValThrGlyLe 615
2998 CTGACACATCAAAATATTGAAACCTTTGACTTACACTTATATAAAA 3047
615 uAlaGlyAspArgThrLysAspPheHisPheGluIleGluLeuLysAsn 632
3048 AGCGGATATACAGACG.....CCACTTAAAGAG 3076
632 snLysGlnGluLeuLeuSerGlnThrValLysThrAspLysThrAsnLe 648
3077 CGAAATTCCTTTACAGCA.....CCAGATACGATATT 3111
649 GluPhe...LysAspGlyLysAlaThr.....IleAsnLe 659
3112 GAATTACCAAAAGATGGCAAGAAACGATCTTGTGTTTGAACCTT 3161
659 uLysHisGlyGlu...SerLeuThrLeuGlnGlyLeuProGlyLys 675
3162 AAAACCGGAAATATGTTCTTACAGAAACCTTACGCCAGAGATATC 3211
675 eTyrLeuValLysGluThrAsp.....SerGluGlyTyrLysVal 688
3212 AGGGATTAAAGAACCAATCAATTAATAATTCGTGAAGATGTTGAGTC 3261
689 LysValAsnSerGlnGluValAlaAsnAlaThrValSer.....Lys 702
3262 ACGATAGATGGGAAAGATGACGATGTTTAATTTCTGAGAGAAAGA 3311
702 sThrGlyIleThrSerAspGluThrLeuAlaPheGluAsnAsnLysGluP 719
3312 TAAATCAATTAATCTTAGACCTAG.....AACCAAGCAAGGTTG 3352
719 roValValProThrGlyValAspGlnLysIleAsnGlyTyrLeuAlaLeu 735
3353 CTTTACTGAACTGTGGCATAGAGCGCTTGTGTTTACTGATATAGCG 3402
736 Ile.....ValIleAlaGlyIleSerLeuGlyLe 745
3403 ATTAGTACATTCGTGATAGCGGGGTATTATCTTATT 3441
seq_name: /cgnl_7/plodata/1/pna/us090_comb.seq:us-09-070-927-45
seq_documentation_block:
; Sequence 45, Application us/09070927
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; APPLICANT: Patrick J. Dillon
; APPLICANT: Steven C. Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and
; NUMBER OF SEQUENCES: 962
; CORRESPONDENCE ADDRESS:

```



```

10112 .....AAATAGTGAACCTGTGAGTATCGA..... 10138
496 Alalleglutyrseryleuthrcluthrclutheularglaalathrgl 512
10139 ATGGTGAACATCGA.....CCATTAGCTACGACACTCT 10176
512 nleuAlailetyrtyrpherthraspseralagluleuaspyspysl 529
10177 TGGCGAAATTATGATTACT.....GTGATATA..... 10207
529 eulysaspyrthsglypneuglyaspmethsnaspserrthleualaval 545
10208 .....TATGGGAAA..... 10216
546 Alalyslleuvalglutyralagl.....Aspsersnpropr 559
10217 .....ATTCACTATGCGCAAAATAATTGAGAAAATGCGCC 10254
559 oglnleuthraspleuasphepelleproasnasnulystyrglns 576
10255 AGATGAGCA..... 10264
576 erleuileglythrglthrphisproulaspuleuvalaspillearg 592
10264 ..... 10264
593 Metgluaspyspysgluvalilleprovalthrhisnleuthrleuar 609
10265 .....CTGACACATCAAAATTAATTGAA 10287
609 glys.....ThrvAlthrnglyleuAlaglyAsparthrlyasp 623
10288 ACCGTTTGACTTACAGTATAAAAAAGCGATATCAGACG..... 10330
623 hehispehguillegluleulysasnulysglnleuLeuSergln 639
10331 .....CCACTTAAAGAGCGAAATTCGCTTAAACAGCA... 10363
640 ThrVallystrhaspyspysThrAsnleuGluphe..LysaspGlyLysAl 655
10364 .....CCAGATACGGATATTGAATTACCAAAAGATGCCAAGA 10401
655 arhr.....IleasnleuLysHISgLyLyu..Serleut 666
10402 AACGGATACCTTTGTTTGAACCTTAAACAGGGAATATGTTCTAA 10451
666 hrleuGlnGlyLeuProGluGlyTyrSeryrleuValLysGluThrAsp 682
10452 CAGAAACCTTTACCGCCAGAGATATCAGGGGTAAAGAACCAATCGAA 10501
683 .....SerglueGlyTyrLysValLysValasnSerglnGluValAl 696
10502 TTAAATAATGCTGAAGATGCTTCAGTCAGATAGTGGGAAAAAGTAGC 10551
696 aasnAlaThrValSer.....LysThrGlyLleThrseraspGlu 710
10552 AGATGTTTATTTCTGAGAGAGAAATATCAATTAATTACCTTACAGCTA 10601
710 hrleuAlaPheGluAsnAsnLysGluProValValProThrGlyValasp 726
10602 CG.....AACCAAGCAAAAGGTTCTTACCTGAAACGTGGGCATA 10642
727 GlnLysLleasnGlyTyrleuAlaLeuile.....Valilleagl 740
10643 GGAAGCTTGTGGTTTACTTGATAGCAGATTAGTACATTGATGATAGCGG 10692
740 YlaserleuGlylle 745
10693 TGTATTATCTTTATT 10708
seq_name: /cgnl_7/plodata/1/pna/US090.COMB.seq:US-09-070-927A-45
seq_documentation_block:

```

```

; Sequence 45, Application US/09070927A
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
;          Patrick J. Dillon
;          Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15614 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-070-927A-45

alignment_scores:
Quality: 196.00 Length: 706
Ratio: 0.596 Gaps: 42
Percent Similarity: 46.601 Percent Identity: 22.946

alignment_block:
US-09-494-297-2 x US-09-070-927A-45 ..
Align seg 1/1 to: US-09-070-927A-45 from: 1 to: 15614

96 ValAsnleuGluGlySer.....ArgSeryrGlnValLyrGysPheAs 110
110 nleuLysLysAlaPheProleuGlySerAspSerSeryrValLysLysTrpT 127
9112 .....TTGGGTAAAGCGCAAGAAATTCAAATTGATT 9142
127 YrlyslYsHISaspGlyLleSerThrLysPheGluasp..... 139
9143 ATCAAGTAGCT.....ATCAACAAGAGTCAGAAACTTCAAACTGAT 9186
140 .....TyrAlaMetSerProAlrGlyLeThrGlyAspGluLeuasnGlnly 154
9187 TTTTGTATCAAAATGAATGCTGCGACAAACGTTTCAGCCATTAGCCACGCG 9236

```

```

154 sLeuArgAlaValMetTyrAsnGlyHisProGlnAsnAlaAsnGlyIleM 171
    :: ||| ::|
9237 CCCGAAAAAGTTGATTTT.....GGGGTTC 9262
171 etGluGlyLeuGluProLeuAsnAlaIleArgValThrGlnAlaVal 187
    ::||| ||| ::|
9263 CTGCGGAAAAAGCACCT.....GGCGGAGTTAAACGTGAAAAAATTC 9306
188 TrpIlyTyrSerAspAsnAlaProIleSerAsnProAspGluSerPheLy 204
    ||| ||| ||| ||| ||| ||| ||| |||
9307 TGGCAGAGTATGATCAAGACCGACCAAGTCGCCAGTAAATGTGATTTA 9356
204 sArgGluSerGluSerAsnLeuValSerThrSer..... 215
9357 TGAATTATGTGAAAGCAGTAACTGACACAGCCAACTGGCAAACTGGCT 9406
216 .....GlnLeuSerLeuMetArgGlnAlaLeuLysGlnLeuIleAspPro 230
9407 ATATTAAATTTCAAAACAGAAATGATACCAGCAATAGTTGGAGCGC 9456
231 AsnLeuAlaThrLysMetProLysGlnValProAspAspPheGln.... 245
9457 AAAAATGTAACCCAACTTCCAAAACCGCGGATGAAAGCTATCAAGAAGT 9506
246 LeuSerIlePheGluSerGluAspLysGlyAspLysTyrAsnLysGlyT 262
9507 TCTTGGGCTTCCCAATACAAATCAAGACAGCACTTTCAT.....T 9550
262 yrgIAsnLeuLeuSerGlyLeuValProThrLysProProThrPro 278
9551 ATCAAAACAAACCGTGAATTAGCA...GTTCCTGTTACAGTCAAGAAAAA 9597
279 GlyAspProMetProProAsnGlnProGlnThrThrSerVal.... 293
9598 ATCGACGATTAAGTCTTGGAAAAACGAGCAAGTCAAGCCATTAGATT 9647
294 ....LeuIleArgLysTyrAlaIleGlyAspTyrSerLysLeuLeuGlu 309
9648 AAAAGTAAATCAAAATTTCTTCAGGTGAG.....AAAACTTAGTGG 9691
309 LysAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheGlnAlaArg 325
9692 GAGCGCTCTTGAATGTAGTGTAATAATGTT.....CAAAACACA 9732
326 ValPheSerSerAsnAsp.....IleGlyIleuArgIleGlu 338
9733 TTATGTGCAATAAAGATGATGCTATTCCTTGCAAAAGATGTGCGCT 9782
338 uSerAspGly...ThrTyrThrLeuThrGluLeuAsnSerProAlaGlyT 354
9783 ACATAAAGGGGACGCTATACATTACGAGTAAACACACCTGCAGAGAC 9832
354 yrsrIleAlaGluProIleThrPheLysValGluAla.....Gly 367
9833 ATGAGTTAGGACAGAAACGACTTGCCAAATGTAGGTAGTGAGCAGAGC 9882
368 LysValTyrThrIleIleAspGlyLysGlnIleGluAsnProAsnLysG 384
9883 AAAGTA.....AGCATCGATGACAGAAAGTGCACACACCAATCAAGT 9926
384 uIleValGluProTyrSerValGluAlaTyrAsnAspPheGluGluPhe 401
9927 TATT.....CCATTGGAATTTGAA.....AATTAATTTCTTCTTGGC 9964
401 erVal.....LeuThrThrGlnAsnTyrAlaLysPheTyrTyr 413
9965 CAATCAGAAATTGAAAAATACACATGCAAAT..... 9996
414 AlaLysAsnLysAsnGlySerSer.GlnValValTyrCysPheAsnAla 430
9997 GCGAAACAAAGTAACTTAGCAGAGCGCACTTTTCGCTGCAAGAA.... 10042
430 sPheLysSerProProAspSerGluAspGlyLysThrMetThrPro 446

```

```

10043 ..... ||| ::|
447 AspPheThrThrGlyGluValLysTyrThrHisIleAlaIleArgAspLe 463
    ::||| ||| ::|
10058 AGTTACCAACTGTGGCAACTCAAAAAACAGATACTACAGGATTCAGCTA 10107
463 uPheLysTyrThrValLysProArgAspThrAspProAspThrPheLeu 480
10108 TTTT..... 10111
480 ySHisIleLysLysValIleGluLysGly...TyrArgGluLysGlyGln 495
10112 .....AAATAGTGAACCTGGTGAATGCA..... 10138
496 AlaIleGlyTyrSerGlyLeuThrGlnIleuArgAlaIleThrG 512
10139 ATGGTGAACAATCAGCA.....CCATTAGCGCTACGACACTCT 10176
512 nLeuAlaIleTyrTyrPheThrAspSerAlaGluLeuAspLysAspLys 529
10177 TGTGGAAATTATGAAATTACT.....GTTGATAAA..... 10207
529 euLysAspTyrHisGlyPheGlyAspMetAsnAspSerThrLeuAlaVal 545
10208 .....TATGGAAA..... 10216
546 AlaLysIleLeuValGluTyrAlaGln.....AspSerAsnProPr 559
10217 .....ATTCACTATGAGCGCAAAAAATATGAAAGAAATGCGCC 10254
559 oGlnLeuThrAspLeuAspPhePheIleProAsnAsnAsnLysTyrGln 576
10255 AGAATGACA..... 10264
576 erLeuIleGlyThrGlnThrHisProGluAspLeuValIleIleArg 592
10264 ..... 10264
593 MetGluAspLysLysGluValIleProValThrHisAsnLeuThrLeuAr 609
10265 .....CTGACACATCAAAATATTGAA 10287
609 gLys.....ThrValThrGlyLeuAlaGlyAspArgThrLysAsp 623
10288 ACCTTTGACTTAACAGTTAATAAAAAAGCCGATATTCAGACG..... 10330
623 heHisPheGluIleGluLeuLysAsnAsnLysGlnGluLeuSerGln 639
10331 .....CCACTTAAAGGAGCGAAATTCGTTTAAACAGA.... 10363
640 ThrValLysThrAspLysThrAsnLeuGluPhe...LysAspGlyLysAl 655
10364 .....CCAGATACGATATTAATTAACCAAAAGATGGCAAGAA 10401
655 arThr.....IleAsnLeuLysHisGlyGlu...SerLeuT 666
10402 AACGGATACTTTTGTGTTTGAACCTTAAACACAGGCAAAATGTCTTCA 10451
666 hrLeuGlnGlyLeuProGluGlyTyrSerTyrLeuValLysGluThrAsp 682
10452 CAGAAACCTTTACGCCGACAGAAAGATATCAGGGGTTTAAAGAAACAATGCA 10501
683 .....SerGluGlyTyrLysValLysValAsnSerGlnGluValAl 696
10502 TTAATTAATTCGTGAAGATGTTCACTCAGATAGATGGGAAAAAGTGC 10551
696 aAsnAlaThrValSer.....LysThrGlyTyrThrSerAspGluT 710
10552 AGAGCTTTAATTTCTGAGAGAGAAAGAAATATCAATTAATCTTAGACGCTTA 10601
710 hrLeuAlaPheGluAsnAsnLysGluProValValProThrGlyValAsp 726
    ||| ||| ||| ||| ||| ||| ||| |||

```

```

10602 CG.....AACCAAGCAAGGTCTTACTTGAAACTGTGGCATA 10642
727 Glnylsileasnglytyrleualaleulle.....Valleialgl 740
10643 GGAGCGTTGGTTTACTTACTGATGATTAGTACATTGCGATAGCGGG 10692
740 YlleSerleuglylle 745
10693 TGTATTCTCTTTATT 10708

seq_name: /cgnl_7/plodata/1/pna/us6004_COMB.seq:us-60-045-649-1048

seq_documentation_block:
: Sequence 1048, Application us/60045649
: GENERAL INFORMATION:
: APPLICANT: Lagace, Robert E.
: APPLICANT: Corely, Neil C.
: APPLICANT: Russo, Frank D.
: APPLICANT: Hann, Amy L.
: TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
: TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
: NUMBER OF SEQUENCES: 1466
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
: STREET: 3174 PORTER DRIVE
: CITY: PALO ALTO
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/60/045,649
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: CERONE, MICHAEL C.
: REGISTRATION NUMBER: 39,132
: REFERENCE/DOCKET NUMBER: PM-0003P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 855-0555
: TELEFAX: (415) 845-4166
: INFORMATION FOR SEQ ID NO: 1048:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7971 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: IMMEDIATE SOURCE:
: CLONE: EF1C1048
: US-60-045-649-1048

alignment_scores:
: Quality: 192.00 Length: 716
: Ratio: 0.598 Gaps: 45
: Percent Similarity: 44.832 Percent Identity: 23.603

alignment_block:
US-09-494-297-2 x US-60-045-649-1048 ..
Align seq 1/1 to: US-60-045-649-1048 from: 1 to: 7971

96 Valasnleuglylser.....ArgSerlyrglnvaltyrcyspheas 110
1672 GTGCGCTAACAGCGCTACATTAATAGTATGAGATTAT..... 1713
110 nleuylslyslalaphneProleuglylserAspSerSerVallylslystrpt 127
|||||.....:|

```

```

1714 .....TTGGGTAAAGGCAAGAAATTCATTCATTCAT 1744
127 YrlslyshisaspGlylleSerThrlysphegluasp..... 139
1745 ATCAAGTAGCT.....ATTCAACAGAGTCAGAAACTTCACAACTGAT 1788
140 .....TyrAlaMetSerProArgIleThrGlyAspGluLeuasnGlnly 154
1789 TTTTGGTATCAATGATGCGGACACAGCTTTCAGCCATTAGCCAGCGC 1838
154 sleuargalavalmetYrAsnnglyHisProGlnAsnAlaAsnGlyIle 171
1839 CCTGAAAAAGTTGATTTT.....GGGGTTC 1864
171 etGlnGlyLeuGluProLeuAsnAlaIleargValThrGlnGluAlaVal 187
1865 CTTCGGGAAAAAGCACCT.....GGCGTGAAGTAAACGTGAAAAAATC 1908
188 TrpTyrTyrSerAspAsnAlaProIleSerAsnProAspGluSer.PheL 204
1909 TGGGAGAGATGATGATCAAGACCCAGACAGTCGCCAGATATATGATTTA 1958
204 ysargyluserglusersnleuvalser.....ThrsGlnLeuSer 218
1959 TGAATTAGTAGAAGCAAGTAAGTACAGACAGCCACTGGCAAACTGGGT 2008
219 leuMetargGlnAlaLeuLysGlnleuIleasp..... 229
2009 ATATTAATTAATTCAAAAACAGAAATGATACCAAGCAATAGTTGGAGCGC 2058
230 .....ProasnleuAlaThrLysMetProLysGlnValProaspa 243
2059 AAAATGTAAACCACTTTCACAAACCGCGATGAAAGCATATCAAGAG. 2107
243 spheGlnLeuSerIlePheGlnSerGlnAspLysGlyAspLysTyrAsn 259
2108 .....TTCCTGGGCTCCGCCCAATACAAATCAAGCAAGCAAGCTTCAT 2152
260 LysGlyTyrGlnAsnLeuLeuSerGlyGlyLeuValProThrLysProp 276
2153 .....TATCAACACACCCGTGAATTAGCA...GTCTCGGTACACTCA 2193
276 othrProGlyAspProPheMetProAsnGlnProGlnThrThrserv 293
2194 AGAAAAATGACGATACTACTTGAAAAACAGACAGACTCAAGCCAT 2243
293 al.....LeuIleargLysTyrAlaIleGlyAspTyrSerLysLeu 306
2244 TAGATTTAAAGTAAATCAAAATCTTCCTCAGGTGAG.....AAAAC 2287
307 leuGlnGlyAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheG1 323
2288 TTACTGGAGCGCTTTGTAATGAGGTAAATATTT.....CA 2328
323 nalaarValalpheSerSerAsnasp.....IleGlyGluArgI 336
2329 AACAAATTAGTGGACAAATAAAGATGTAGCTATTCCTGGCAAAAGATG 2378
336 leGluLeuSerAspGly...ThrTyrThrleuThrGluLeuAsnSerPro 351
2379 TGCCCTACAAAAAAGGGAACGCTATTAATCAAGTAAGTAAGAACACCT 2428
352 AlaglyTyrSerIlealagluProIleThrPheLysValGluAla..... 366
2429 GCAGACATAGATTAGGCAAGAAAAAGCACTTGCAAAATTAAGGTGACTGA 2478
367 .....GlyLysValTyrThrIleIleAspGlyLysGlnIleGluAsnPro 382
2479 GCAAGCAAACTA.....AGCATGATGACACAGAAAGTACACCAACCA 2522
382 snlysgluIlealagluProTyrSerValGluAlaTyrAsnAspPheGlu 398
2523 ATCAAGTTATT.....CCATTTGGAATTTGAA.....AATTAATTTTCT 2560

```

```

399 GluPheSerVal.....LeuThrThrGln..... 406
      ::::: ||| |||
2561 TCCTTGCCAAAGCAGATTAGAAAATACACATGCAAAATGGCAAAACAGT 2610
      ||| |||
407 AspTyrAlaLysPheTyrTyrAla.....LysAsnLysAsnGly 420
      ||| |||
2611 GACCTTACAGAGGAGGACTTTGGCTTGCAGAAAGAAAATGCTCAGAGAA 2660
      ||| |||
420 eSerGlnValValTyrCysPheAsnAlaAspLeuLysSerProProAsp 436
      ||| |||
2661 GTTACCAAAAGTGGCAGACATCAAAAACAGAT..... 2692
      ||| |||
437 SerGluAspGlyGlyLysThrMetThrProAspPheThrThrGlyGluVal 453
      ||| |||
2693 .....ACTACAGCA..... 2701
      ||| |||
453 LysTyrThrHisIleAlaGlyArgAspLeuPheLysTyrThrValLysP 470
      ||| |||
2701 ..... 2701
      ||| |||
470 roArgAspThrAspProAspThrPheLeuLysHisIleLysValIle 486
      ||| |||
2702 .....TTGACCTATTTT...AAATATGT 2722
      ||| |||
487 GluLysGly...TyrArgGlyLysGlyGlnAlaIleGlyTyrSerGlyLe 502
      ||| |||
2723 GACCTGCTGATGATCGA.....ATGGCGCAACATACAGCA.. 2758
      ||| |||
502 uThrGlnThrGlnLeuArgAlaIleThrGlnLeuAlaIleTyrTyrPhe 519
      ||| |||
2759 .....CCATTAGGCTACGACACTCTTGCTGGAAATATGATGATTA 2798
      ||| |||
519 hrAspSerAlaGlnLeuAspLysAspLysLeuLysAspTyrHisGlyPhe 535
      ||| |||
2799 CT.....GTGCATAAA.....TAT 2812
      ||| |||
536 GlyAspMetAsnAspSerThrLeuAlaValAlaLysIleLeuValGluTyr 552
      ||| |||
2813 GGGAAA.....ATTCACCTA 2826
      ||| |||
552 rAlaGln.....AspSerAsnProProGlnLeuThrAspLeuAspP 566
      ||| |||
2827 TGCAGGCAAAATATTGAGAAAATGCCCGCAATGAGCA..... 2866
      ||| |||
566 hePheIleProAsnAsnLysTyrGlnSerLeuIleGlyThrGlnTrp 582
      ||| |||
2866 ..... 2866
      ||| |||
583 HisProGluAspLeuValAspIleIleArgMetGluAspLysGluVal 599
      ||| |||
2866 ..... 2866
      ||| |||
599 IleProValThrHisAsnLeuThrLeuAlaGlyLys.....ThrValT 613
      ||| |||
2867 .....CTGACACATCAAAATTAATTGAACCTTTGACTTACAGATTC 2909
      ||| |||
613 hrGlyLeuAlaGlyAspArgThrLysAspPheHisPheGluIleGlyLeu 629
      ||| |||
2910 ATAAAAAGCGGATATACAGAG.....CCACTT 2938
      ||| |||
630 LysAsnAsnLysGlnGluLeuLeuSerGlnThrValLysThrAspLysTh 646
      ||| |||
2939 AAAGGAGCGAAATTCCTTTAACAGGA.....CCAGATAC 2973
      ||| |||
646 rAsnLeuGluPhe...LysAspGlyLysAlaThr.....I 657
      ||| |||
2974 GGATTTGAATTTACCAAAAGATGGCAAAAGAACGGATCTTTGTTTGTG 3023
      ||| |||
657 leAsnLeuLysHisGlyLys...SerLeuThrLeuGlnGlyLeuProGlu 672
      ||| |||
3024 AAACTTAAACCAAGGAAATATGTCTTAACAGAAACCTTTACGCCAADA 3073

```

```

673 GlyTyrSerTyrLeuValLysGluThrAsp.....SerGluGlyTyr 686
      ||| |||
3074 GGATATCAGGGGCTTAAAGAACCAATCGAATTAATTAATTCGTGAAGATGG 3123
      ||| |||
686 rLysValLysValAsnSerGlnGlnIleValAlaAsnAlaThrValSer.... 701
      ||| |||
3124 TTCACCTCAGATAGATGGGAAAAAGTAGCAGATGTTTAATTTCTGGAG 3173
      ||| |||
702 .....LysThrGlyIleThrSerAspGluThrLeuAlaPheGluAsnAsn 716
      ||| |||
3174 AGAAGAAATATCAATTAATCTTACGCTTACG.....AACCAAGCA 3214
      ||| |||
717 LysGluProValValProThrGlyValAspGlnLysIleAsnGlyTyrLe 733
      ||| |||
3215 AAGGTCCTTACCTGAAACGAGGCGCATAGACGCTGTGTTTACTT 3264
      ||| |||
733 uAlaLeuIle.....ValIleAlaGlyIleSerLeuGlyIle 745
      ||| |||
3265 GATAGCGATTAGTACATTCGTGATAGCGGCTGTATCTCTTATTT 3310
      ||| |||

```

seq_name: /cgnl_7/plodata/1/pna/us6004.COMB.seq:us-60-046-653-1042

seq_documentation_block:

; Sequence 1042, Application US/6004653

; GENERAL INFORMATION:

; APPLICANT: Lagace, Robert E.

; APPLICANT: Corley, Neil C.

; APPLICANT: Russo, Frank D.

; APPLICANT: Hann, Amy L.

; APPLICANT: Heath, Joe D.

; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS

; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF

; NUMBER OF SEQUENCES: 1449

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/60/046,653

; FILING DATE: HEREWITH

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: CERRONE, MICHAEL C.

; REGISTRATION NUMBER: 39,132

; REFERENCE/DOCKET NUMBER: PM-0003-1 P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 855-0555

; TELEFAX: (415) 845-4166

; INFORMATION FOR SEQ ID NO: 1042:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8040 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; IMMEDIATE SOURCE:

; CLONE: EFALC1042

; US-60-046-653-1042

alignment_scores:

; Quality: 192.00

; Ratio: 0.598

; Percent Similarity: 44.832

Length: 716

Gaps: 45

Percent Identity: 23.603

alignment_block:
US-09-494-297-2 x US-60-046-653-1042 ..

Align seg 1/1 to: US-60-046-653-1042 from: 1 to: 8040

```

96 ValAsnLeuGluGlySer.....ArgSerTyrGlnValTyrCysPheAs 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1741 GTGCGCTACAGCGCTACAAATAGTAATGAGATTTAT..... 1782
110 nLeuLysLysAlaPheProLeuGlySerAspSerSerValLysTyrTrp 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1783 .....TTGGTAAAGGCGCAAGAAATTCAAATTCATT 1813
127 TyrLysLysAspGlyLysSerThrLysPheGluAsp..... 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1814 ATCAAGTACGT.....ATTCAAACACAGTCACGAAAACTTCAACCTGAT 1857
140 .....TyrAlaMetSerProArgIleThrGlyAspGluLeuAsnGlnLys 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1858 TTTTGGTATCAATGAAATGTCGAGCAACGTTTCAGCCATTAGCCACGGC 1907
154 sLeuArgAlaValMetTyrAsnGlnHisProGlnAsnAlaAsnGlyIleM 171
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1908 CCTCGAAAGAGTTGATTT.....GGGCTTC 1933
171 etGluGlyLeuGluProLeuAsnAlaIleArgValThrGlnLysAlaVal 187
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1934 CTTCGGGAAAAAGCACCT.....GGCTGAAGTTAAACGTAAAAAATC 1977
188 TrpTyrTyrSerAspAsnAlaProIleSerAsnProAspGluSer.PheL 204
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1978 TGGGACAGATGATCAAGACCGACCAAGCGCCAGATTAATGTGATTTA 2027
204 ysaArgGluSerGluSerAsnLeuValSer.....ThrSerGlnLeuSer 218
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2028 TGAATATTAGTAAAGCAAGTAACTGACACAGCCCACTGGCAAACTGGCT 2077
219 LeuMetArgGlnAlaLeuLysGlnLeuIleAsp..... 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2078 ATATTTAAATTATCAAAACCGAAAAATGATACCACCAATAGTTGGGAGCC 2127
230 .....ProAsnLeuAlaThrLysMetProLysGlnAlaProAspA 243
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2128 AAAAATGTAACCACTTCCAAAAACCGCGATGATGAAGCTATCAAGAAG. 2176
243 sprPheGlnLeuSerIlePheGluSerGluAspLysGlyAspLysTyrAsn 259
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2177 .....TTCCTGGCTTCCCAATATCAACCAATCAAGACCAAGCTTTCAT 2221
260 LysGlyTyrGlnAsnLeuLeuSerGlyLysLeuValProThrLysProPr 276
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2222 .....TATCAACAACACCCGTGATTTAGCA...GTCTCTGCTTACATCA 2252
276 oThrProGlyAspProMetProProAsnGlnProGlnThrThrServ 293
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2263 AGAAAAATGACGATCTACTTGGAAAAACACGAGACAGTTCACAGCAT 2312
293 al.....LeuIleArgLysTyrAlaIleGlyAspTyrSerLysLeu 306
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2313 TAGATTTTAAAGTATCAAAAATTTCTTCAGGTGAG.....AAAAAC 2356
307 LeuGluLysAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheG 323
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2357 TTAGTGGAGACCGCTTGTGAATTGAGTGTAAAAATGTT.....CA 2397
323 nAlaArgValPheSerSerAsnAsp.....IleGlyGluArgI 336
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2398 AACAACTATTAGTGCACAATAAAGTAGAGTATTCCTTGCACAAAGATG 2447
336 leuGluLeuSerAspGly...ThrTyrThrLeuThrGluLeuAsnSerPro 351
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2448 TGGCCCTACAAAAAGGGGAACGCTATACATTAACTGAAGTAAAAAGCACCT 2497

```

```

352 AlaGlyTyrSerIleAlaGluProIleThrPheLysValGluAla..... 366
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2498 GCAAGCATGATGATTAGGCAAGAAACGACTTGGCAATTTAGGTAGCTGA 2547
367 ....GlyLysValTyrThrIleLeuAspGlyLysGlnIleGluAsnPro 382
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2548 GCAGGCGAAAGTA.....AGCATCGATGGACAAAGAGTACCACCAACAA 2591
382 snLysGluIleValGluProTyrSerValGluAlaTyrAsnAspPheGlu 398
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2592 ATCAAGTTATT.....CCATTGGAAATTGAA.....AATTAATTTTCT 2629
399 GluPheSerVal.....LeuThrThrGln..... 406
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2630 TCTTTGCCAATACAGAAATTAGAAATACACATGCAAAATGCAACAAGT 2679
407 .AsnTyrAlaLysPheTyrTyrAla.....LysAsnLysAsnGly 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2680 GAACCTTAGCAGAGCGCACTTTGGTTGCACAAAGAAAAAATGCTCAAGAA 2729
420 eLysGlnValValTyrCysPheAsnAlaAspLeuLysSerProProAsp 436
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2730 GTTACCAAACTGTGGCACTCAAAAACAGAT..... 2761
437 SerGluAspGlyLysThrMetThrProAspPheThrThrGlyLysVal 453
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2762 .....ACTACAGA..... 2770
453 LysTyrThrHisIleAlaGlyArgAspLeuPheLysTyrThrValLysP 470
2770 ..... 2770
470 roArgAspThrAspProAspPheLeuLysHisIleLysLysValIle 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2771 .....TTGAGCTATTTT...AAATATTCT 2791
487 GluLysGly...TyrArgGluLysGlyGlnAlaIleGlyTyrSerGlyLe 502
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2792 GAACCTGTGATGATGCA.....ATGGTGCACACACAGAA... 2827
502 uThrGluThrGlnLeuArgAlaIleThrGlnLeuAlaIleTyrTyrPheT 519
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2828 .....CCATTAGCTACGACACTCTTCTGGAATATGAAATTTA 2867
519 hrAspSerAlaGluLeuAspLysAspLysLeuLysAspTyrHisGlyPhe 535
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2868 CT.....GTGATTA.....TAT 2881
536 GlyAspMetAsnAspSerThrLeuAlaValAlaLysIleLeuValGlyTyr 552
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2882 GGGAAA.....ATTCACTA 2895
552 rAlaGln.....AspSerAsnProProGlnLeuThrAspLeuAsp 566
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2896 TGCAGCGCAAAATATTGAGAAAAATGCGCCAGATGACAA..... 2935
566 hePheIleProAsnAsnLysTyrGlnSerLeuIleGlyThrGlnTrp 582
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2935 ..... 2935
583 HisProGluAspLeuValAspIleIleArgMetGluAspLysGluVal 599
2935 ..... 2935
599 IleProValThrHisAsnLeuThrLeuAlaGlyLys.....ThrValT 613
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2936 .....CTGACACATCAAAATATTGAAAACCTTTGACTTAACAGTTC 2978
613 hrGlyLeuAlaGlyAspArgThrLysAspPheHisPheGluIleGluLeu 629
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2979 ATAAAAAAGCGGATATTCAGACG.....CCACTT 3007
630 LysAsnAsnLysGlnIleLeuLeuSerGlnThrValLysThrAspLysTh 646

```



```

||||| ||| :||| :|||
3008 AAGGACCGCAAAATTCGGTTTAACAGA.....CCAGATAC 3042
646 rAsnLeugluPhe..LysAspGlyLysAlaThr.....I 657
||||| ||| :||| |||
3043 GGATATTGAATTACAAAAGATGCCAAGAAACGATACTTGTGTTTTG 3092
657 IeAsnLeuYshISgLYGU...SerLeuThrLeuInGLyLeuProGLU 672
||||| ||| :||| |||
3093 AAAACTTAAAAACGAGGAATATGTTCTTAACAAGAACCTTTAGCCAGAA 3142
673 GLYTyrSerTYrLeuValLySGluThrASP.....SerGLUGlyty 686
||||| ||| :||| |||
3143 GGNATATCAGGGGTTAAAAAGAACCAATGCATTAATATTTGTTAAATG 3192
686 rLVsValLySvalAsnSerInGLUValAlaAsnAlaThrValSer.... 701
||||| ||| :||| |||
3193 TTCAGTCAGCATGATGGGGAAGAAAGTAGCACGATTTAATTTCTGGAG 3242
702 ....LysThrGLylleTherSerAspGLUthrLeuAlaPheGLUasnaSn 716
3243 AGAAGATATATCAAATTACTTTAGACGTTAGC.....AACCAAGA 3283
717 LysGLUProValProThrGLyValAspInLysILeasnGLyTyLe 733
||||| ||| :||| |||
3284 AAGGTTCTTTTACCtTGAAACGTGGCGCATAGACGCTTGtGGTTTACCT 3333
733 uAlaLeuLeu.....ValIIeAlaGLylleSerLeuGLyTyLe 745
||||| ||| :||| |||
3334 GATAGCGATTAGTACATTTCGTATAGCGGGGtTTATCTTTATT 3379

seq_name: /cgn1_7/ptodata/1/pna/us6006.COMB.seq:US-60-068-217-919
seq_documentation_block:
Sequence 919, Application US/60068217
GENERAL INFORMATION:
APPLICANT: Lagace, Robert E.
APPLICANT: Corley, Neil C.
APPLICANT: Russo, Frank D.
APPLICANT: Hann, Amy L.
APPLICANT: Heath, Joe D.
APPLICANT: Finney, Gregory L.
TITLE OF INVENTION: NOCLOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
NUMBER OF SEQUENCES: 1239
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: word perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/068, 217
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PM-0003-5 P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-4166
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 919:
SEQUENCE CHARACTERISTICS:
LENGTH: 11907 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

```

```

;
; MOLECULE TYPE: genomic DNA
;
; IMMEDIATE SOURCE:
; CLONE: EFAC919
; US-60-068-217-919

alignment_scores:
      Quality: 192.00      Length: 716
      Ratio: 0.598      Gaps: 45
      Percent Similarity: 44.832      Percent Identity: 23.603

Alignment block:
US-09-494-297-2 x US-60-068-217-919 ..

Align seg 1/1 to: US-60-068-217-919 from: 1 to: 11907

96 ValAsnLeuGlueLysSer.....ArgSerTyrGlnValTyrCysPheAs 110
|||:||| |||:||| |||:||| |||:|||
5478 GTCCTCCCTACAGCGCCCTACATTAAAGTAAAGACAGTTAT..... 5519
110 nLeuLysLysAlaPheProLeuGlnLysSerAspSerValLysLysTrp 127
5520 .....TTGGGTAAAGCGCAAGAAATTCAAATTCAT 5550
127 yLysLysHisAspGlyLleSerThrLysPheGluAsp..... 139
|||:||| |||:||| |||:||| |||:|||
5551 ATCAAGTACGT.....ATTCAACAGAGATCAGACAAACTCAACCTGAT 5594
140 .....TyrAlaMetSerProArgLleThrGlyAspGluLeuAsnGlnLy 154
||| |||:||| ||| ||| ||| ||| |||
5595 TTTTGGTATCAATGATGATGTCGCGACACACTTTCAGCCATTAGCAGCGC 5644
154 sLeuArgAlaValMetCysAsnGlyHisProGlnAsnAlaAsnGlyLeu 171
||| ||| ||| ||| ||| ||| ||| |||
5645 CCTCGTAAAGATTGATTT.....GGGGTTC 5670
171 eGlnGluLysLeuGlnProLeuAsnAlaIleArgValThrGlnGluAlaVal 187
|||:||| ||| ||| |||:||| |||:||| |||:|||
5671 CTTCGGGAAAAGCAGCT....GGCGTGAAGTTAAACGTGAAAAATTC 5714
188 TrpTyrTyrSerAspAsnAlaProLleSerAsnProAspGluSer.PheL 204
||| |||:||| ||| ||| |||:||| ||| ||| |||
5715 TGGGAAGAGATGATCAAGACCGACAGATCGGCCAGATATATGATTTA 5764
204 yAsArgLysGluSerAsnLeuValSer.....ThrSerGlnLeuSer 218
||| |||:||| ||| ||| ||| ||| ||| |||
5765 TCGAAATTAGTAAAGCAACAGTATCTACACAGCCAACTGGCAAACTGGT 5814
219 LeuMetArgGlnAlaLeuLysGlnLeuIleAsp..... 229
|||:||| ||| ||| ||| ||| ||| ||| |||
5815 ATATTAAATTATCAAAACCAAGAAATGATCCAGCAATAGTTGGGAGCC 5864
230 .....ProAsnLeuAlaThrLysMetProLysGlnValProAspA 243
|||:||| ||| ||| ||| ||| ||| ||| |||
5865 AAAAATGTAAACCACTTCCAAACCGCGAGAAAGTATCAAGAAG. 5913
243 sPheGlnLeuSerLlePheGlnSerGlnAspLysGlyAspLysTyrAsn 259
|||:||| |||:||| |||:||| |||:||| |||:|||
5914 ....TTCTGGGCTTCCGCCAATACAAATCAAGACCAAGCTTTCAT 5958
260 LysGlyTyrGlnAsnLeuLeuSerGlyLysValProThrLysProPr 276
|||:||| |||:||| |||:||| |||:||| |||:|||
5959 .....TATCAACCAACCCGTGAATTAGCA..GTTCCTGGTTACAGTCA 5999
276 oThrProGlyAspProProMetProProAsnGlnProGlnThrThrSery 293
||| ||| ||| ||| ||| ||| ||| ||| |||
6000 AGAAAAATTCAGAGATCACTTGAAGAAAAACAGCAACAGTTCACAGCAT 6049
293 al.....LeuLleArgLysTyrAlaIleGlyAspTyrSerLysLeu 306
|||:||| |||:||| |||:||| |||:||| |||:|||
6050 TGGATTAAAGTAATCAAAAATTTCTCCCTCAGGTGAG.....AAAAC 6093

```



```

Align seg 1/1 to: US-09-663-779-1301 from: 1 to: 10627

208 GluSerAsnLeuValSerThrSerGlnLeuSerLeuMetArgGlnAlaLe 224
|||||
3989 GAAATGAAATGGTGGATACGAGGAATGTAGAGATA..... 4024
224 uLysGlnLeuLeuAspProAsnLeuAlaThrLysMetProLysGlnValP 241
|||||
4025 .....ACAAAATAGATTAAGATAGTA 4046
241 roAspAspPheGlnLeuSerLlePheGlnSerGluAsp...LysGly... 255
|||||
4047 AACGTCATTAGAAATGTGTATTGAGTACGCTGATTTAAAGGAAAA 4096
256 .....AspLysThrAsnLysGlyTyrrGlnAsnLe 265
|||||
4097 GTAGTGCAAAAGTACGACGATAAAGAAAGAAACCAACGTTTCAGA 4146
265 uLeuSerGlyGly.....LeuValProThrLysProProThrProG 279
|||||
4147 TTATATCTATGGAAAGTATGAGTTAGTAAGAAACACCGGCGAGAT 4196
279 LysAspProPheProAsnGlnProGlnThrThrSerValLeuIle 295
|||||
4197 ACAAACCACTAGAAAGCCAAATTCATTCGAAATGAAAGGTAAGATA 4246
296 ArgLysThrAlaIleGlyAspThrSerLysLeuGlnGlyAlaThrLe 312
|||||
4247 ACACGATTACAAATTGACTGTAGAAATGAATTAAGTATACGAGAAATGT 4296
312 uGlnLeuThr.....GlyAspAsnValAsnSerPheGlnAlaArgV 326
|||||
4297 AGAAATTACAAAAGTAGATAAAGAAATTAAGATGCTTACGTCATGAC 4346
326 alPheSerSerAsnAspLleGlyLysArgLe..... 336
|||||
4347 TCTTGGAAATTCAGATGAACGACAGCAAGTACGCTAAATTAACGACA 4396
337 .....GlnLeuSerAspGlyThrTyrrTh 344
|||||
4397 GATTAATAAGACACAAAGCAAGTACTAATTTATCAGTCGACATACAA 4446
344 rLeuThrGlnLeuAsnSerProAlaGlyTyrr...SerLleAlaGluProI 360
|||||
4447 GTTAGTGAAGTAAAGACCAAAAGATTAACCAATAGTAGATCGA 4496
360 lThrPheLysValGlnAlaGlyLysValTyrrThle.....Ile 373
|||||
4497 TTACTTTCCAAATTTGAAAAGGCAATGACAAAATCTCTTGCTTTAACGTA 4546
374 AspLylLysGlnIleGlnAsnProAsnLysGlnLeu.....ValGluPr 388
|||||
4547 GAAACGAAATGTTAGACAAAGGAAATGTGAGTAACAAAGTAGATAA 4596
388 cTyrrSerValGlnAlaTyrrAsnAspPheGlnGluLysSerValLeuThrT 405
|||||
4597 AGATAGTCAAAAAGCATTA..... 4615
405 hrcGlnAsnTyrrAlaLysPheTyrrAlaLysAsnLysAsnGlySerSer 421
|||||
4616 .....GCA 4618
422 GlnuValValTyrrCysPheAsnAlaAspLeuLysSerProProAspSerG 438
|||||
4619 AGTGTACTCTTC.....GAAGTACACAGA 4641
438 uAspGlyGlyLysThrMetThrProAspPheThrThrGlyGlnValLysT 455
|||||
4642 CGAACCAAGGCAAGTAGTAACA...GAAGTACGACGAGATTAAGAAAGCA 4688
455 yTrpHisIleAlaGlyArgAspLeuPheLysTyrrThrValLysProArg 471
|||||
4689 AAGCAAAATCTCAGACTATCTGTAGGAAGTAC..... 4723
472 AspThrAspProAspThrPheLeuLysHisIleLysLysValIleGlyL 488
|||||
4724 .....AAATAGTAGAGAA 4737
488 s.....GlyTyrrArgGlyLysGlyGlnAlaIleGluTyrr.... 499
|||||
4738 AAAAAGCTTACAGGTTTACAAAAAAGTACAGACGACGATCATTCAGAA 4787
500 .....SerGlyLeuThrGlnuThr...GlnLeuArgAlaAlaThrGlnLeu 513
|||||
4788 TTAAGAGGCTATGACAGAGCTCTTATCATTAAGAGAAAGAAAGCAAG 4837
514 AlaIleTyrrThrPheThrAspSerAlaGlnLeuAspLysAspLysLeu 530
|||||
4838 TTACACAAAGCTTTAGTAGAAATCACAAAAGTAAAGTAAAGTATCAAAA 4887
530 sAspTyrrHisGly.....PheGlyAspMetAsnAspSerThrLeuAla 545
|||||
4888 AGTATTAGAAAGCGTACTCTTCGAAAGTACAGATGAAAAAGCAAAAGTAG 4937
545 alAlaLysIleLeuValGlnuTyrrAlaGlnAspSerAsnProProGlnLeu 561
|||||
4938 TAACAAAAGTACACACAGAT.....AAAGAAAGCAAAAGCAAGATT 4978
562 ThrAspLeuAspPhePheIleProAsnAsnLysTyrrGlnSerLeuI 578
|||||
4979 GCTGATTATATCAGTA.....CGAAATATATAG...CTAGT 5010
578 eGlyThrGlnThrHisPro.....GluAspLeuValAspLleLeuArg 593
|||||
5011 AGAGGTAGAGCGTTCACCGATTAATAAACTAGCGCAAAACCATACAT 5060
593 eGluAspLysLysGlnuValIleProValThrHisAsnLeuThrLeuArg 609
|||||
5061 TTGAAATCAAAAAA.....GCTATGACAGAAAGCTTATCTCACTAAA 5101
610 .....LysThrValThrGlyLeuAl 616
|||||
5102 GTAGAGATGAATGACTAGTAGACAAAGCTCAGTAGAAATCACAAAAGTGA 5151
616 agLysAspArgThrLys.....AspPheHisPheGlnIleGluLeu 630
|||||
5152 TAAAGATAGTCAAAAGATATTAGAAGACGTAGTCTGCAAGTCAA... 5197
630 ysaAsnAsnLysGlnGluLeuLeuSerGlnThrValLysThrAspLysThr 646
|||||
5198 ....GACGACCAAGGAAAGTACTTAACGCAAGTAAACACGAT..... 5236
647 AsnLeuGlnPheLysAspGlyLysAlaThrIleAsn..... 658
|||||
5237 .....AAAATGGTAAGCAAAAATCTCAGACTATCTGATAG 5274
659 .....LeuLysHisGlyGlnSerLeuThr...LeuGlnGlyLeuP 671
|||||
5275 AAAGTACAAATTTAGTAGAAAGAGGCTTACAGGCTACAAAAAACTAA 5324
671 roGlnGlyTyrrSerTyrrLeuValLysGlnuThrAspSerGlnGlyTyrrTh 687
|||||
5325 CAGAGCCAGTATCATTCGAATTTAAAAAGGCTATGACAGAGAGTCTTATCA 5374
688 ValLysValAsnSerGlnGlnuValAlaAsnAlaThrValSerLysThrG 704
|||||
5375 TTAAGACTAGAGATGAACAGTTAGCAAAAGGTTCACTAGAAATACAAA 5424
704 yLleThrSerAsp.....GluThrLeuAlaPheGlnu... 714
|||||
5425 AGTAGATTAAGATAGTCAAAAAGTATTAGAAAGCGGTGCTTCGCAATAC 5474
715 ...AsnAsnLysGlnuProValValProThrGlyValAspGlnLysIleAsn 730
|||||
5475 AAGCGACAAAGCAAGTAGTA...ACAGAGTAAAAACAGATAAAAAT 5521

```



```

344 hrleuthrGluleuanserProAlaGlyTyrSerIleAlaGluProIle 360
1236 CGCTGACGAAACAAAGCTCCAGAGGGCATGTAAAGCAAAAAGACT 1305
361 ThrPheLysValGluAlaGlyLysValTyrThrIle...IleAspGlyLys 376
1306 ACTGGGAAATCAACAGATCGCTTCATGATGCTACGTAACCATTTGAGAAA 1355
376 S.....GlnIleGluAsnProA 382
1356 AACAGTCACACTCTCCATGATAGCATGCTGACATGTGAAAAATCCTT 1405
382 snLysGluIleValGluProTyrSerValGluAlaTyrAsnAspPheLys 398
1406 TTGTTGAAGTT.....CCTGTACGATACGTAACATGCGATCGAACAGG 1449
399 GluPheSerValIleuThrThrGlnAsnTyrAlaLysPheTyrTyrAlaLys 415
1450 ACGGCAAAAGATTAATCTTAAGGACAGCATTTCCCTACAGAAAAA 1499
415 sAsnLysAsnGlySerSerGlnValTyrCysPheAsnAlaAspLeuL 432
1500 AGAAGCAATGTGCTATCTATCAG..... 1521
432 ySerProPheAspSerGlu.....AspGlyGlyLysThrMetThr 445
1522 ....CCAATTGACAGCCAAACAAACGAATGAAAAAGTCTGCCAGTTT 1566
446 ProAspPheThrThrGlyGluValLysTyrThrHisIleAlaGlyArgAs 462
1567 GATTCACACACACCGTAAATATCGACTCGTGAACACGCTGCCTGC 1616
462 pleuPheLysTyrThrValLysProArgAspThrAspProAspThrPheL 479
1617 CGGATAT.....GATACCTCGCCGGAATATATG 1645
479 euLysHisIleLysLysValIleGluLysGlyTyrArgGluLysGlyLys 495
1646 AATTCCAATCGATAAATAT.....GGAATA 1671
496 AlaIleGluTyrSerGlyLeuThrGlnThrGlnLeuArgAlaIleThrGly 512
1672 ATCATTT...TACACGGGAAAAAATACCGAGATGACAAATAATGTATGAC 1718
512 nLeuAlaIleTyrTyrPheThrAspSerAlaGluLeuAspLysAspLysL 529
1719 GCTCAGTCATCAAAATCGACTAAAGCCTTGATCTACGGAACCAAAA 1768
529 euLysAspTyrHisGlyPheGlyAspMetAsnAspSerThrIleuAlaVal 545
1769 AAGAAGAC.....AACGACAGACATTAAGAAGA 1797
546 AlaLysIleLeuValGluTyrAlaGlnAspSerAsnProProGlnLeuTh 562
1798 GCAAAATTCAGACTCAG.....GACCCAGAAATG.. 1827
562 rasPheLysPhePheIleProAsnAsnAsnLysTyrGlnSerLeuIleG 579
1828 .GACTTAGAA..... 1836
579 LyrThrGlnTyrHisProGluAspLeuValAspIleIleArgMetGluAsp 595
1837 .....TCGCCAAAGAT..... 1848
596 LysLysGluValIleProValThrHisAsnLeuThrLeuArgLysThrVa 612
1848 ..... 1848
612 ThrGlyLeuAlaGlyAspArgThrLysAspPheHisPheGluIleGluL 629
1849 .....GGACAAGAAACGATACCTTTCTATTGCA..... 1878
629 euLysAsnAsnLysGlnIleuLeuLeuSerGlnThrValLysThrAspLys 645

```

```

1878 ..... 1878
646 ThrAsnLeuGluPheLysAspGlyLysAlaThrIleAsnLeuLysHisGly 662
1879 .....AATTAAACCTCG 1892
662 yGlu...SerLeuThrLeuGlnGlyLeuProGluGlyTyrTyrSerTyrLeu 677
1893 AACTTATCGCTGACGAAACTTTTACACAGAAAGATACCAAGGCTCTAA 1942
678 .....ValLysGluThrAspSerGluGlyTyrLysValLysValAsn 691
1943 AAGACCCGATTACTTACTTATACAGAAAGATGGGTCAATTCAATGATGAT 1992
692 SerGlnGluValAlaAsnAlaThrValSerLysThrGlyIleThrSerAs 708
1993 GGACAAAGAT.....CATGAATCTGTTCTGTACACAGAGCAAGCAAAAAACA 2036
708 pGlu...ThrIleuAlaPheGluAsnAsn...LysGluProValValProT 723
2037 CCAGATTCCTTAGACATCAGCAATCAGCAAAAGTACCATTAACCTGAA 2086
723 hrGlyValAspGlnLysLysHisGlyTyrLeuAlaIleuIleValIleAla 739
2087 CGGAGGAATGGCCGTTTAGAATCTAT.....CTAGTAGGATGATTT 2130
740 GlyIleSerLeuGlyIleTyrP 746
2131 GGTGTGCGTTTCTATTG 2151

seq_name: /cgn1_7/prodata/1/pna/us6006_COMB.seq:us-60-068-186-748
seq_documentation_block:
: Sequence 748, Application us/60068186
: GENERAL INFORMATION:
: APPLICANT: LAGACE, ROBERT E.
: APPLICANT: CORLEY, NEIL C.
: APPLICANT: RUSO, FRANK D.
: APPLICANT: HANN, AMY L.
: APPLICANT: HEATH, JOE D.
: APPLICANT: THORNTON, MICHAEL
: APPLICANT: FINNEY, GREGORY L.
: TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECIUM
: TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USBS THEREOF
: NUMBER OF SEQUENCES: 1053
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
: STREET: 3174 PORTER DRIVE
: CITY: PALO ALTO
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/60/068,186
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: CERRONE, MICHAEL C.
: REGISTRATION NUMBER: 39,132
: REFERENCE/DOCKET NUMBER: PM-0005-3P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 855-0555
: TELEFAX: (650) 845-4166
: INFORMATION FOR SEQ ID NO: 748:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 18768 base pairs
: TYPE: nucleic acid

```

STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: genomic DNA
 IMMEDIATE SOURCE:
 CLONE: EFM1c748
 US-60-068-186-748

alignment_scores:
 Quality: 180.00 Length: 741
 Ratio: 0.347 Gaps: 41
 Percent Similarity: 44.399 Percent Identity: 20.783

alignment_block:
 us-09-494-297-2 x US-60-068-186-748 ..

Align seg 1/1 to: US-60-068-186-748 from: 1 to: 18768

72 TrpTyr.....GlyTyrGluSerTyrValArgLysH1 82
 |||||
 12459 TGGTACACTTTCAGTCAAGAGTGGGACAGACCTTAC...AAGGTCA 12505
 82 s.ProTyrTyrLysGlnPheArgValAlaHisAspLeuArgValAsnLeu 98
 |||||
 12506 CTCCA.....TCTATTTCCTA 12522
 99 GluLysSer.....ArgSerTyrGlnValTyrCysPheAsnLeuLysLys 113
 |||||
 12523 GAAGGAATACCATCAAGACCAATCAGATCTAT..... 12555
 113 sAlaPheProLeuGluSerAspSerSerValLysLysTyrLysLysH 130
 |||||
 12556TTAGAAAGACCAAGAAATCCAAATCCATTACCAAGTGA 12595
 130 tsAspGlyLysSerThrLysPheGluAspTyr..... 140
 |||||
 12596 GA.....ATCCAAACAGAAATGAGAGCTTCATCCAAATTTCTGTAT 12639
 141AlaMetSerProArgLysThrGlyAspGluLe 151
 |||||
 12640 CAATGACGAGCAGACACATTTCCAGCCAAACATTCATACCAATGATTT 12689
 151 uAsnGlnLysLeuArgAlaValMetTyrAsnGlyHisProGlnAsnAla 168
 |||||
 12690 AGCTGAA.....T 12697
 168 snGlyLysMetGluGlyLeuGlnProLeuAsnAlaIleArgValThrGln 184
 |||||
 12698 TCGGTATACCATCTCTGTAAGCTCCGAGTCACTTCACATC..... 12741
 185 GluAlaValTyrTyrTyrSerAspAsnAlaProLysSerAsnProAspG 201
 |||||
 12742 AAAAAGTTATGGGAAGATTGACACACATCTACGTGTCGACAGATCA 12791
 201 u.....SerPheLysArgGluSerGluSerAsnLeuValSerThrSerGln 217
 |||||
 12792 AGTACTTTTTCGATTCACGGA.....CATACGACCAATG 12829
 217 euSerLeuMetArgLysAlaLeuLysGlnLeuIleAspProAsnLeuAla 233
 |||||
 12830 CTGCAGCTTGGAAAAACGGATATTCGAATCAATTAACCGATTAAGAT 12879
 234 Thr.....LysMetProLysGlnVal 240
 |||||
 12880 ACAACAATACGTGGGACGTGACACATTGACAAATATATCTGCAAAATAG 12929
 240 LProAspAspPheGln.....LeuSerIlePheGluSerGluAspLysG 255
 |||||
 12930 CGGAGAAAGTTATCAAGAGATATATACCTACCAATACAAATATATCAAG 12979
 255 LysAspLysTyrAsnLysGlyTyrGlnAsnLeuLeuSerGlyGlyLeuVal 271
 |||||
 12980 GTCAAGCATTCAGT.....TACCAACAATC..... 13005

272 ProThrLysProProThrProGly.....AspProPr 282
 |||||
 13006 ...AAAGATATACCTGTACCAAGATACGATTCACAAATATAGTGCAT 13052
 282 oMetProProAsnGlnProGlnThrThrSerVal.....LeuIleA 296
 |||||
 13053 GACATGGAAAAATACATAACATTCACACCGTTAAACCTGAAAAATACGA 13102
 296 rGlyTyrAlaIleGlyAspTyrSerLysLeuGluGlyAlaThrLeu 312
 |||||
 13103 AAAATTCCTCTGCGAGTGA.....AAGATTCATCGGCTGTCTTC 13146
 313 GlnLeuThrGlyAspAsnValAsnSerPheGlnAla...ArgValPhe 328
 |||||
 13147 AAATTAACAGAGATTCATTTGATCTTACTAATACATATCATGCGACGG 13196
 328 rSerAsnAspIleGlyGluArgIleGluLeu...SerAspGlyThrTyr 344
 |||||
 13197 AACCTATTCTCTCCAGAAAAATGCATAATTCAGAAAAATGACCTATA 13246
 344 hLeuThrGlnLeuAsnSerProAlaGlyTyrSerIleAlaGluProIle 360
 |||||
 13247 CGTGACAGAAACAAAGCTCCAGACAGGCAATTAACGAAAAAGACT 13296
 361 ThrPheLysValGluAlaGlyLysValTyrThrIle...IleAspGly 376
 |||||
 13297 ACTTGGGAAATCAAGATCCGCTTCGTGATGTACGTATACCTGATGTA 13346
 376 s.....GlnIleGluAsnPro 382
 |||||
 13347 AACAGTCTACTCTCCGATGATACATCCAGTTGACATTAATGAAATCTT 13396
 382 snLysGluIleValGluProTyrSerValGluAlaTyrAsnAspPheGlu 398
 |||||
 13397 TTGTGAAGTT.....CCTGACAGTACGTAAAGTATGCGAATGCAAGG 13440
 399 GluPheSerValLeuThrGlnAsnTyrAlaLysPheTyrTyrAla 415
 |||||
 13441 ACGGACAAAGATTAATCTTAAGAGAGCAGCATTTCCCTACGAAAAA 13490
 415 sAsnLysAsnGlySerSerGlnValValTyrCysPheAsnAlaAspLeu 432
 |||||
 13491 AGAAGCAAAATGTACTATACG..... 13512
 432 ySerProProAspSerGlu.....AspGlyGlyLysThrMetThr 445
 |||||
 13513CCAATTGACAGCCAAACAAAGATGAAGAAAGGCTTGGCCAGTTT 13557
 446 ProAspPheThrThrGlyGluValLysTyrThrHisIleAlaGlyArg 462
 |||||
 13558 GATTCACCTACACCTGTGTAATATACGAGTGTGAACAGCTGCTCTGC 13607
 462 pLeuPheLysTyrThrValLysProArgAspThrAspProAspThrPhe 479
 |||||
 13608 CGGATAT.....GATACCTCCCGGGAATATATG 13636
 479 euLysHisIleLysValIleGluLysGlyTyrArgGluLysGlyGln 495
 |||||
 13637 AATTCCAAATCGATTAATAT.....GGAAAA 13662
 496 AlaIleGluTyrSerGlyLeuThrGluThrGlnLeuArgAlaAlaThrG 512
 |||||
 13663 ATCATTT...TACACGGGAAAAAATACGAGATGACAAATATATATGAC 13709
 512 nLeuAlaIleTyrTyrPheThrAspSerAlaGluLeuAspLysAspLys 529
 |||||
 13710 GCTCAGCATCAAAATGACATAAAGGCTTGTGATCAACGTTACACAAAA 13759
 529 euLysAspTyrHisGlyPheGlyAspMetAsnAspSerThrLeuAlaVal 545
 |||||
 13760 AAGAAGAC.....AACGACAGACATTTAAAGA 13788

```

546 AATLALYSLILEUVALIETRYTALACILNAPSERSAPROPOGILNETH 562
|||||: : : : : |||||: : : : :
13789 GCAGAAATTCAGACTGCAG. . . . . GCACAGGAATG. . 13818
562 rASpLeuSAPhEPhE1leProASnASnLysTYrGlnSerLeuIleG 579
|||||: : : : :
13819 .GACTTAGAATG. . . . . 13830
579 lYThGlNTrPhISPrOGlUnSPlEUValASpIleIleArgMetGlnASP 595
|||||: : : : :
13831 . . . . . CCAAAAGAT. . . . . 13839
596 LysLysGluValIleProValThrHisASnLeuThrLeuArgLysThVa 612
13839 . . . . . 13839
612 lYhRgLYleuAlaGlyASpArgTYhRAsPPhEHisPheGluIleGluL 629
|||||: : : : : ||| |||||
13840 . . . . . GGACAAAGAAACGATACCTTCTTATTCGA. . . . 13869
629 eULysASnASnLysGlnGluLeuLeuSerGlnThrValLysThrAsPlys 645
13869 . . . . . 13869
646 ThrASnLeuGluPhElyASpGlyLysAlaThrIleASnLeuLysHisG1 662
13870 . . . . . AATTTAAACCTGG 13883
662 yGlu. . .SerLeuThrLeuGlnGlyLeuProGluGlyTyrSerTYrLeu. 677
| : : : : |||||: : : : :
13884 AACTTATACGCGTGCAGCAACTTTACACGCAAGATACCAAGTCTCAA 13933
678 . . . . . ValLysGluThrASpSerGluGlyTYrLysValLysValASn 691
||| : : : ||| : : : : :
13934 AAGAGCCACTTACTACTAGTTATACAGCAAGATGGCTCAATTCAGTGGAT 13983
692 SerGlnGluValAlaASnAlaThrValSerLysThrGlyIleThrSerAS 708
|||||: : : : : ||| : : : : :
13984 GSACAGAGAT. . . . . CATGAATCTGTTCTGTGCACGAGAGCCAAAAACCA 14027
708 pGlu. . .ThrIleAlaPheGluASnASn. . .LysGluProValValProT 723
||| : : : ||| : : : : :
14028 CCAGATTCTTTAGACATCAGCAATCAGCAAAATACCATCTACTGAAA 14077
723 hGcLYleASpGlnLysIleASnGlyTYrLeuAlaLeuIleValIleAla 739
||||| : : : : : ||| : : : : :
14078 CGGAGAGGATTCGCCGTTTAGAATCTAT. . . . . CTAGTAGGATGATT 14121
740 GlyIleSerLeuGlyIleTrp 746
||| : : : : : |||||
14122 GGTGTGCGCTTTCTATATTGG 14142

seq_name: /cgnl_7/prodata1//pna/us6005_COMB.seq:us-60-050-444-744

seq_documentation_block:
  : Sequence 744, Application US/60050444
  : GENERAL INFORMATION:
  : APPLICANT: APPLICANT: lagace, Robert E.
  : APPLICANT: Corley, Neil C.
  : APPLICANT: Russo, Frank D.
  : APPLICANT: Hann, Amy L.
  : APPLICANT: Health, Joe D.
  : TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
  : NUMBER OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USUS THEREOF
  : NUMBER OF SEQUENCES: 1063
  : CORRESPONDENCE ADDRESS:
  : ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
  : STREET: 3174 PORTER DRIVE
  : CITY: PALO ALTO
  : STATE: CALIFORNIA
  : COUNTRY: USA
  : ZIP: 94304

  : COMPUTER READABLE FORM:

```

1 MEDIUM TYPE: Floppy disk
 2 COMPUTER: IBM PC compatible
 3 OPERATING SYSTEM: PC-DOS/MS-DOS
 4 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2.
 5 CURRENT APPLICATION DATA:
 6 APPLICATION NUMBER: US/60/050,444
 7 FILING DATE:
 8 CLASSIFICATION:
 9 ATTORNEY/AGENT INFORMATION:
 10 NAME: CERRONE, MICHAEL C.
 11 REGISTRATION NUMBER: 39,132
 12 REFERENCE/DOCKET NUMBER: PM-0003-2P
 13 TELECOMMUNICATION INFORMATION:
 14 TELEPHONE: (415) 855-0555
 15 TELEFAX: (415) 845-4166
 16 INFORMATION FOR SEO ID NO: 744:
 17 SEQUENCE CHARACTERISTICS:
 18 LENGTH: 18786 base pairs
 19 TYPE: nucleic acid
 20 STRANDEDNESS: single
 21 TOPOLOGY: linear
 22 MOLECULE TYPE: cDNA
 23 IMMEDIATE SOURCE:
 24 CLONE: EFM1c744
 25 US-60-050-444-744

```
alignment_scores: 100.00
```

Quality:	180.00	Length:	14
Ratio:	0.547	Gaps:	41
Percent Similarity:	44.399	Percent Identity:	20.783

alignment_block:
US-09-494-297-2 x US-60-050-444-744

Align seg 1/1 to: US-60-050-444-744 from: 1 to: 18786

72 TPTPTT.....GlyTyrGluSerTyrValArgGlnI 82

12459 TGGTACACTTTTCAGTCAAGAGTGTGGGCAAGTCTTAC...AACGGTCA 12503

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible]

113 2) b) b) Definiere die Begriffe "Kultur" und "Werte". (20 Punkte)

TTTAGGAAAGACCAAGAAATCCAAATCCATTACCAAGTGA 12595

130 ISASPGVILESerThrLysPheGluASPTyr..... 140

12596 GA...ATCCAACAGAAATGAGGACTTCATCCAAATTCTGTAT 12639

141AlaMetSerProArgIleThrGlyaspGluLe 151

12640 CAAATGACGGCAGGACAACTTTCAGCCAAACATTGTATACCAATGAATT 12689

151 uasnglnlysleuargalaValmettyrasnglyhisproglinasnalaA 168

12090 AGC1GAA.....T 12091

[illegible][illegible]


```

seq documentation block:
: Sequence 28, Application US/09596002
: GENERAL INFORMATION:
: APPLICANT: Lagace, Robert, E.
: APPLICANT: Patterson, Chandra
: APPLICANT: Berg, Kim, L.
: TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
: FILE REFERENCE: PM-0008-4 US
: CURRENT APPLICATION NUMBER: US/09/596, 002
: CURRENT FILING DATE: 2000-06-16
: PRIOR APPLICATION NUMBER: 60/140,121
: PRIOR FILING DATE: 1999-06-18
: NUMBER OF SEQ ID NOS: 41
: SOFTWARE: PERL Program
: SEQ ID NO 28
: LENGTH: 49617
: TYPE: DNA
: ORGANISM: M. catarrhalis
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte template ID No: 28
: PUBLICATION INFORMATION:
: US-09-596-002-28

alignment_scores:
      Quality: 177.00      Length: 798
      Ratio: 0.515      Gaps: 40
      Percent Similarity: 43.108      Percent Identity: 20.551

alignment block:
US-09-494-297-2 x US-09-596-002-28 ..

Align seg 1/1 to: US-09-596-002-28 from: 1 to: 49617

      84 TyrTyrLysGlnPheArgValAlaHisAspLeuArgValAsnLeuGlnGlu 100
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
14887 TATTTGATTAATTCCTCCCAATATCCGATCTG.....CACCTGAGAAA 14930
      100 ySerArgSerTyrGlnValTyrCysPheAsnLeuLysLysAla..... 114
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14931 CAGCGAG.....CATGTTTGTGATGCTCAAAAAACCAATATACA 14968
      115 .....PheProLeuGlySerAspSerValLysLys..... 125
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
14969 TCAAAATATATGTTATGTCATGTCATCATCCCTGCCAAAAACCAACC 15018
      126 .....TprTyrLysLysHisAspGlyLysSerThrLys.....PheG1 138
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15019 TACATGATTTTCACACAGAACAAACATCAAAAAACCAACAGCGCA 15068
      138 uAspTyrAla.....MetSerProArgIleThr.... 147
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15069 TGATTATCAAAACATGCTTTTGGCTATATGAGCTAAGAGAGCGGAGCC 15118
      148 .....GlyAspGlyLeuAsnGlnLysLeuArgAlaValMet 159
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15119 TAAATTAATAAAAGTCAGACACCCAGCGCAAGAACCCGTCATCATTT 15168
      160 TyrAsnGly.....HisProGlnAsnAlaAsnGlyIleMe 171
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15169 TTCACACACCTACTTATTTTATCATGTGAGATGCCAGACCCCATCT 15218
      171 tGluGlyLeuGluProLeuAsnAlaIleArgValThrGlnGluAlaVal 188
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15219 GCCA.....AAGCGGGGTAAATTGACATATGAGGGCATTT 15253
      188 rPtyrTyrSerAspAsnAlaPro.....IleSerAsnProAsp 200
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15254 GGTGTATCTACGATCGATGCAAAAAAGCCCATTTTATGATTAACAGAC 15303
      201 Glu.....SerPheLysArgGlySerGlySerGlySer 210
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

15304 GATAAGTAGGCACTATTTTAACTCAACGAGAAATCAATGAAGCGCA 15353
210 nLeuValSerThrSerGlnLeuSerLeu..... 219
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15354 TTTGGTAGTCAGACACACATTTATCTAACAGCTTTAAATATAACACA 15403
220 .....MetArgGlnAla 223
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15404 CCCCCGCCACTTATAGCGTCGACTTTGATCAAAATACCTTAAGGCCAAA 15453
224 LeuLysGlnLeuIleAspProAsnLeuAlaThrLysMetProLysGlnVal 240
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15454 TTGCTCTTATTTGACAAACCAACAGCAACAGCCGATGGCGCTTATAT 15503
240 lProAspAspPhe.....GlnLeuSerIlePheG 250
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15504 CAGAAGTAGCTTTGATATACGCAAAAAAGTCATATGACCCGATGCTATG 15553
250 lSerGlnAspLys.....GlyAspLysTyrAsnLysGlyTyrGlnAsn 264
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15554 AGATTGACGCCAAGATTATATGGAACCGCTTACTGGCAGACCAAAATCT 15603
265 LeuLeuSerGlyGlyLeu.....ValProTh 273
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15604 TTGATTTGATGATTAACACCAATACCGACCTTTTGTAAAGAGCTGTCTC 15653
273 rLysProProThrProGlyAspProPromeProAsnGlnProGlnT 290
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15654 CAAAAAGCCAAATCCCAACCAACCA.....GACCCCAAC..... 15687
290 hrThrSerValLeuIleArgLysTyrAlaIleGlyAspTyrSerLysLeu 306
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15688 .....TCAGATACG 15696
307 LeuGlnGlyAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheG1 323
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15697 CTAGAGGCGGGGTTTATGTCGAGTCGGGCGCATGAGCTG..... 15735
323 nAlaArgValPheSerSerAsnAspIle.....G 333
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15736 GCGGGTAATTTTATTCATGACACGCAACTTTTGTGCTTTGGGTG 15784
333 LysGlnArgIleGlnLeuSerAspGlyThrTyrThrLeuThrGlnLeuAsn 349
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15785 GCAACGCGACAAAAACGACGACCTGTGCCACAAAAACGGTGATTTT 15834
350 SerProAlaGlyTyrSerIleAlaGluProIleThrPheLysValGluAl 366
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15835 AGT...ACAGATTTGAA.....AAACCCAGACACGACTTTGGTGC 15875
366 aGlyLysValTyrThrIleIleAspGlyLys.....GlnIleG 379
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15876 TGAAGAGATGTGTGACATTTATGACGGTAAAGTAAATGATGAGTCA 15925
379 lAsnProAsnLysGlnIleValGluProTyrSerValGluAlaTyrAsn 395
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15926 ATAATCAAAATTGAAGATGAACCTGCTGTCAATTAATGAATATAT 15975
396 AspPhe.....GluGluPheSerVal 402
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15976 GAATATATTTATGACGACGCCAACAAACATTCACCAAAAAAATTAACGC 16025
402 lLeuThrThrGlnAsnTyrAla.....LysPheTyrT 413
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
16026 CAGGTCGCAAAAAACCTGCTTATTTTGTACAGCATGATGAAGTTTAT 16075
413 yr.....AlaLys.....AsnLysAsnGly 419
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
16076 TTAATGTAACTATTATGACTTATACGCCAAAGAAAGCAAAACCTTGCT 16125
420 SerSerGlnValValTyrCysPheAsnAlaAspLeuLysSerProProAs 436
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
16126 GTCTCCCAAGATACGACGACCAATRAAGATTTTGGCTAAATACCCAGA 16175

```


15219	GCCA.....	AAGCGGGTAAATTGACATGTAGGGCAATT	15253
188	rPtyrTyrSeraspasnalaPro.....	IleSeraspProasp	200
15254	GGTTGATCTGCACCGATGCCAATAAACGCCCATTTT	TAGATAAACAGAC	15303
201	GLU.....	SerPheIysatrgLusergLusSer	210
15304	GATAAAGTAGGCATTTATTTTAACTCAACACAGAAATATCAAAAGAGGCGA		15353
210	nleuValSerThrSerGlnLeuSerLeu.....		219
15354	TTTGGTAGTGCAGCACACATTTATCTAAACGCTTTAAATTAACACA	15403	
220	MetargGlnala	223
15404	CCCCGGCCACTTATAGCGTGGACCTTGATCAATAATACCTTAAGAGCAAA	15453	
224	LeuLysGlnLeuIleaspProasnleualathrLysMetProLysGlnIva	240	
15454	TTGTCTTATTATGACAAACCAACAGCAACAGCGGATGGCGCTTAAAT	15503	
240	LProaspAspPhe.....	GlnLeuSerIlelepp	250
15504	CAGAGCTCAGTTTGATACCGACAAAAGGTCAATAGACCGATGTCTATG	15553	
250	LusergLusAspLys.....	GlyAspLysTyrAsnLysGlyTyrGlnAsn	264
15554	AGATTGACGCCCAAGATTATATGCGAACCGCTTACGTGGCACACCCAAACT	15603	
265	LeuLeuSerGlyGlyLeu.....	ValProth	273
15604	TTGATTGATGATACAAACCAACATACCGACCTTTGTTAAAGACCTGTCTGC	15653	
273	rLysProProthrProGlyAspProProMetProProasnGlnProGlnT	290	
15654	CAAAAAGCCAAATCCCAACAAACCCA.....	GACCCCAAC	15687
290	hrThrSerValleuIleargLysTyrAlaIleGlyAspTyrSerLysLeu	306	
15688	TCAATACG	15696
307	LeuGlnGlyAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheG	323	
15697	CTAGAAGCGGGCTTTTATGTGTAGCGGGCGATGAGCTG.....	15735	
323	nalaIargValPheSerSerAsnAspIle.....	G	333
15736	.CGGGTAAATTTTATCAATAGAACAGCAACCACTTTGTGTCTTTGGC	15784	
333	LylValargIleGlnLeuSeraspGlyThrTyrThrLeuThrGlnLeuAsn	349	
15785	GCAAAAGAGACAAAACGACCGACCTGTGCACAAAACGGTGATATTT	15834	
350	SerProAlaGlyTyrSerIlealaGlnProIleThrPheLysValGluAl	366	
15835	AGT.....ACAGGATTTGAA.....	AAACCCAGCACCAAGCTTGTGGCAA	15875
366	acLysValIlyrThrIleIleaspGlyLys.....	GlnIleG	379
15876	TGAAGAGATTGGTACCATTTATTTGACGGTAAAGAAATTAATGATGAAGCA	15925	
379	LuanPProasnLysGlnIleValGlnLupProTyrSerValGlnIatIyrAsn	395	
15926	ATTAATCAATTTGAAGATGAACATGTCCCTGTCTGACGTAAATTAAGAAATATAT	15975	
396	AspPhe.....	GlnGluPheSerIva	402
15976	GATATATATTTATGACAGACCAACAACAATTCACCAAAAAAATTAAGC	16025	
402	IleuThrThrGlnAsnTyrAla.....	LysPheTyrT	413
16026	CAGCGCGCAAAAAAACCCGTATTATTTGGTCAGCATGATTAAGATTATTT	16075	

413	yr.....	AlaLys.....	AsnLysasnGly	419
	::			
16076	TTATAGTACTATTATGACTTATCATGCACCAAGCAACAAAGCTTGCT			16125
420	SerSerGlnValValTyrCysPheasnAlaAspLeuLysSerProAs			436
		:::		
16126	GTTCTCCCAAGATACCAAGCACCATAAGAAATATTTGGCTAAATACCCAGA			16175
436	pSerGlnAspGlyGlyLysTyrThrMetThrProAspPheThr.....			450
16176	TGCC.....	AAAGTAAGCAGACAGACAATAAAGTTACCAAAATCG		16213
450			450
16214	TTCTACAAAGCCAAAGATAAGCCGTATACCCGATTCATGCCAAAAAGC			16263
451	GlyGluValLysTyrThrHisIleAlaGlyAsr		461
16264	TATGACCACATCAGTTTGGTGAAGTATGTAATGATACAAAGGCA			16313
461	GasPheLeuPheLysTyrThrValLysProAsThrAspProAsPThr			478
16314	CCCAACAGCAGTATTTGTGTGCAAGCGGTCAAGCGATGTCATACTC			16363
478	heLeuLys.....			His
				His
16364	AGCTGCCAGTCGAGTAATTCACCTATAATGCTTTGGCAGGCTAC			16413
482	IleLysLysValIleGluLysGlyTyrArgGlyLysGlyGlnAlaIleG			498
16414	CTGACCCCAAAAAAACAAGGTTATACCAAGATGAGATCAGCATCA			16463
498	uTyrSerGlnLeuThrGlnGlnLeuArgAlaIaIaThrGlnLeuAlaI			515
16464	GCAAAAAGGCTTAAAGAT.....			16482
515	LeuTyrTyrPheThr.....	AspSerAlaGluLeu		524
16483	..TATATATTGACCAAGACTTATCCCAAGATGAGCATGAGCATGAC			16530
525	AspLysAspLysLeu.....	LysAspTyrHis		533
16531	GATGACGATAGTTGGACCGCATCTGATGATTCACAAGATGATAAATACA			16580
533	scGlyPheGlyAspMetAsnAspSerThrIleuAlaValAlaLysIleu			550
16581	TGGCGATGATGATTTG.....	ATTGCAT		16603
550	AlaGluTyrAlaGlnAspSerAsnProGlnLeuThrAspLeuAspPhe			566
16604	CTGATGATTCACACAGATGATGACGCGAGATGGCGATGACATTCATAT			16653
567	PheIleProAsnAsnAsnLysTyrGlnSerLeuIleGlyTyrGlnTyrP			583
16654	TTG.....	GGGATGATGTCAGATGATGACGCCGACGCAAGATGTATCA		16697
583	sProGluAspLeuValAspIleLeuArgMetLysPylsLysGluValI			600
16698	TGCAGGTAAAT.....	ATTTCGCCCTTAATTTGAAAACAAAATACT		16735
600	LeuProVal.....	ThrHisAsnIleuThrLeu.....		608
16736	TGCCATTATATAGCTACTCATGAAAAAACCTTTGCCCTAATGTATAA			16785
609	ArgLysThrValThrGlyLe		615
16786	AATAAGGCTAAGTTGATGATAACTTTGACACCAACAGACGCTAATCGTAA			16835
615	uAlaGlyAspArgThrLysAspPheHisPheIuLysGluLeuLysAsnA			632
16836	ATTAAACGTATGAGAGAGTATATCTGCTCTTATATATTC.....	AAAAATG		16879

```
632 snlvsnglnleuLeuSerGlnThrValIshrsAspLysThrAsnLeu 648
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
16880 GCAAAATTGTCACACAGATTTCACGCCCAAGCCATGCGCAACTAT 16929
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
649 GluPheLysAspGlyLysAlaThrIleAsnLeuLysHisGlyLysSerIe 665
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
16930 CGTGAAGAACTGGT.....ACCAACCAAGTGGC..... 16959
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
665 umrleuGlnGlyLeuProGlnGlyLysSerIleValLysGluThr 682
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
16960 .....GGTCTTATTACAAACATCAACATGATTTG 16987
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
682 spserGlnGlyLysValLysValAsnSerGlnGluAla 696
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
16988 ATGTTAAGGCGCAATTTTGGCACAATGGCGAAGCTGGCA 17031
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

seq_name: /cgn1_7/ptdata/1/pna/US096D_COMB.seq:US-09-663-779-1383

seq_documentation_block:
; Sequence 1383, Application US/09663779
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Shukla, Hridayabhiranjan
; TITLE OF INVENTION: BACILLUS THURINGIENSIS CHROMOSOMAL GENOME SEQUENCES AND USE
; TITLE OF INVENTION: THERMOF
; FILE REFERENCE: 38-21(5137)B
; CURRENT APPLICATION NUMBER: US/09/663,779
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 1999-09-17
; NUMBER OF SEQ ID NOS: 8283
; SEQ ID NO 1383
; LENGTH: 19992
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-663-779-1383

alignment_scores:
Quality: 171.50 Length: 901
Ratio: 0.420 Gaps: 53
Percent Similarity: 45.283 Percent Identity: 19.423

alignment_block:
US-09-494-297-2 x US-09-663-779-1383 ..
Align seq 1/1 to: US-09-663-779-1383 from: 1 to: 19992
9 LysLeuAsnThrIleuAsnThrGlnArgValLeuSerLysAsn..... 22
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
5800 AAAATTGCTACTGACAAATCCAGAAACAAAGCTATCAACGACCTCGTGA 5849
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
23 .....SerLysArgPheThrValThr...LeuValGly 33
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
5850 GAATACAAATTCAGTCCAAAGCAAAATTAACAACAATTTTAACCGGTG 5899
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
33 alPheLeuMetIlePheAlaLeuValThrSerMetValGlyAlaLysThr 49
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
5900 TAAAGTTG.....ACGATTAAGACGGAACCA... 5928
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
50 ValPheGlyLeuValGluSerSerThrProAsnAlaIleAsnProAsp 66
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
5929 .....TTTATAGAGATTAACCGTCAAGTACAGATTCCTT...GC 5966
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
66 rSerSerGlnLysArgPheLysGlyLysGluSerLys...ValArgGly 82
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
5967 CAATATGAGTTTACATGGAACCTTTAAATTCATGATGAAAAATG 6016
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
82 LsProLysThr.....LysGlnPheArgValAlaHisAsp 93
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6017 GAGTACTATATTTTGTGATCTCTTAACAAATTTAAGATTTACAAATCA 6066
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```

```
94 LeuArgValAsnLeuGlnGlySerArgSerLysGlnValLysCysPheAs 110
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6067 ATTAACAGTCCTTATATGATGTAAGCAAAACCAATGTAATTTTAC 6116
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
110 nLeuLysLysAlaPheProLeuGlySerLysSerValLysLysTrp 127
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6117 TGTACAAAA..... 6126
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
127 LysLysLysHisAspGly...IleSerThrLysPheGluAspTyrAlaMet 142
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6127 .....GATGAAAAAGTCACATGATGATCAACGATTTATGTTGA 6165
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
143 Ser...ProArgLysThrGlyAsp.....GluLeuAsnGln 153
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6166 GACATTCACAGATGTTCTGCTGTAACCTACATTAAGACAGAAATTAATA 6215
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
153 n..... 153
6216 AGCTGAATTAAGTTACAAACACACAGAAATTCCTTCCCGATTAAAG 6265
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
154 ..LysLeuArgAlaValMetLysAsnGlyHisProGln..... 165
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6266 ATTAAGATGTTCTATATACAGTTGACCTTTAAGCTTATGTAACACAGCT 6315
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
166 ..AsnAlaAsnGlyIleMetGluGlyLeuGluProLeuAsnAlaIleArg 181
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6316 ACGAATTAAGAAAGGTTTACCTGAT.....AGACCAATTAAT..... 6351
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
181 gValThrGlnGluAlaValTrp..... 188
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6352 ....ACAAATGAATTAATTAATGGAAGTAGAATGACAAACCAAGACAC 6397
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
189 .....TyrTyrSerAspAsnAlaPro.....IleSerAsn 198
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6398 CCCTTAAGAAACGCTTTTAAAGATTAATCCACAGCTACACAGCTTA 6447
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
199 ProAspGluSerPheLys..... 204
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6448 AATTAAGATTTCTAATTAAGTTATTATTATTAAGTTGATGTTAATGGAA 6497
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
205 .....ArgGluSerGluSerAsnLeuValSer 214
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6498 TGTAAACCGGTGGTGCACAGACCGACACAGATTAACAATATTATTTCAT 6547
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
214 hr.....SerGlnLeuSerLeu...MetArgGlnAlaLeuLysGlnLeu 227
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6548 CAGATGCTCAAAATTTGAGATTCCTTCAAGATCTATCAATTAAGTGA 6597
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
228 IleAspProAsnLeuAlaThrLysMetProLysGlnValProAspAsp 244
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6598 TATCAAAATTCGAATTCGCCACAAAATTAAGTGAATTAATAAGTTT 6647
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
244 eGlnLeuSerIlePheGlnSerGluAspLysGlyAspLysTyrAsnLys 261
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6648 CCGAAATTAACGTTACGATTAACGATGATTAATCAAAACCAACCAACGCA 6697
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
261 LysTyrGlnAsnLeuLeuSerGlyLysValProThrLys..... 274
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6698 GCTCTACTGTAAAGCTTTCTGCTGTACACACTTAATTAACCAAGTAA 6747
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
275 .....ProProThrPro.....GlyAs 280
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6748 TATGATCCAAAGACCAACCAATTAAGTGAATTAATTAATTAATGATG 6797
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
280 ProProMet.....P 284
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6798 TCAAGAGAGATCAAAAAACGATGCGATTTTAAAGATTTTTCACG 6847
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
284 roProAsnGlnProGlnThrThrSerValLeuIleArgLysTyrAlaIle 300
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
6848 ATACACAGCAAGTATGCAATTCATTTGTGTAAAAAACGCTCATAT 6897
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
301 GlyAspTyrSerLysLeuLeuGlnGlyAlaThrLeuGlnLeuThrGlyAs 317
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```

```

6898 AATGAAAAAGGACGCTT.....GTAACAGGAGA 6926
      ::::: :::: |||
317 pAsnValAsnSerPheGlnAlaArgValPheSerSerAsnAspIleGly 334
      ||| ||| :::::
6927 CGCTGTTAATTAATAC.....ACTGTAAGTAAACAAAGAATGAT 6967
      ::::: |||
334 LuArgIleGluLeuSerAspGlyThrTyThrLeuThrGluLeuAsnSer 350
      ::::: |||
6968 TCGATTACAGCTTAATGAT.....GATATTAAATAGT 6999
      ::::: |||
351 ProAlaGlyTyrSerIleAlaGluProIleThrPheLysValGlu..A1 366
      ||| ||| :::::
7000 .....GCTTATGTGATTACGATAAACCAGGCCAAC 7031
      ::::: |||
366 aGlyLysValTyrThrIleIleAspGlyLysGlnIleGluAsnProAsnL 383
      ::::: |||
7032 TAAATAAAGTT.....ATAGAAAGTGGAAA...GTAAAAAT 7065
      ::::: |||
383 ySgLIleValGluProTyrSerValGluAlaTyrAsnAsp..... 396
      ::::: |||
7066 .....AAAGTTACGCTGATTAATGATTCAAGTAA 7095
      ::::: |||
397 .....PheGluGluPheSerValLeuThrThrGlnAsnTy 408
      ||| ||| :::::
7096 GAAATGACAGCTACCTTCACAGCACAATATTAATTAATCTAATTAATA 7145
      ::::: |||
408 rAlaLysPheTyrTyrrAlaLysAsnLysAsnGlySerSerGlnValAlT 425
      ||| ||| :::::
7146 AGCTGAACAATATTAAGACAAACAACTGGCAATTAATAGTAA 7195
      ::::: |||
425 yRcysPheAsnAlaAspLeuLysSerPro.....ProAspSerGluAsp 439
      ||| ||| :::::
7196 ATAATTAACACTATCCATTAAACACGCGATTATTACGACACCTTGAT 7245
      ::::: |||
440 gLyLysThrMetThrProAspPheThrThrGlyValLys..... 454
      ||| ||| :::::
7246 CATGCTGATACAAATTAAGATTAAGAACTA...GAAATTAAGACCG 7292
      ::::: |||
455 ...TyThrHisIleAlaGlyArgAspLeuPheLysTyThrValLysP 470
      ||| ||| :::::
7293 AAATTAACCTTCACAGCTGGAGCTGAC..... 7320
      ::::: |||
470 rOArgAspThrAspProAspThrPheLeuLysHisIleLysValIle 486
      ::::: |||
7321 .....TAAGTTTAGATGATACA 7338
      ::::: |||
487 gLuLysGlyTyrArgGluLysGlyGlnAlaIleGluTyrSerGlyLeuTh 503
      ::::: |||
7339 GATTAAGGCTCAAAATTAATCTATAGTACGATCAGTCTAATATGAC 7388
      ::::: |||
503 rGluThrGlnLeuArgAlaIleThrGlnLeuAlaIleTyrTyrrPheThra 520
      ||| ||| :::::
7389 AAAGACA.....TTAGTCGTAAATAATACGACAG 7417
      ::::: |||
520 sPSeRAlaGluLeuAspLysAspLysLeuLys.....AspLysr 532
      ||| ||| :::::
7418 AC.....TTTGATTATACAAAGCTAGAAAGTGTAAACCTTCATTT 7458
      ::::: |||
533 HisGlyPheGlyAspMetAsn.....AspSerThrLeuAlaIaVa 545
      ||| ||| :::::
7459 AAAATATACAGTACCTTCTTGATAGATACGATTCATTAATCCACATC 7508
      ::::: |||
545 lAlaLysIleLeuValGluTyrAlaGlnAspSerAsnProGlnLeuT 562
      ||| ||| :::::
7509 AAAATAAGTTACAGCTAATTCGATTCGATAC..... 7542
      ::::: |||
562 hrAspLeuAspPhePheIleProAsnAsnAsnLysTyrrGlnSer..... 576
      ||| ||| :::::
7543 .....TTCACAAAGCGAATGTTATTAATACGCTTCTTAATAC 7581
      ::::: |||
577 .....LeuIleGlyThrGlnTrpHisProG1 585
      ::::: |||

```

```

7582 GCCCAAGCAAGAAATTAATCTGATTAATAGGCTTAATATAATATGT 7631
585 uAspLeu.....ValAspIleIleArgMetGluAspLysL 597
      ::::: |||
7632 TGAGTTAAAGATCCTTATGTTAATAGAGTCATA.....CAGATTAAC 7675
      ::::: |||
597 ySgLIleValIleProValThrHisAsnLeuThrLeuArgLysThrValHr 613
      ::::: |||
7676 AAAAGTTAGTGCCT.....GATTCATTGAAGTGGCGGATATGATTTTA 7719
      ::::: |||
614 gLyLeuAlaGlyAspArg.....ThrLysAs 622
      ::::: |||
7720 AATGGAATTCCTGATTAATATGACAAACCTGTAATGCTGTACCACTGAGA 7769
      ::::: |||
622 pPheHisPheGluIleGluLeuLysAsnAsnLysGlnLeu..... 636
      ::::: |||
7770 ATATGCACTTGACAGCACTTACAGAGAACATTAACACCTTAAGGTTT 7819
      ::::: |||
637 ..LeuSerGlnThrValLys..... 642
      ::::: |||
7820 ATTTCAAACAAATCAATTAATACCTTACTATATTATCTTAACAAAC 7869
      ::::: |||
643 .....ThrAspLysThrAsnLeuGluPheL 651
      ||| ||| :::::
7870 CTGATGTGAACTTATCCAGGCAACCTACAAAAAGAGGAGATTAA 7919
      ::::: |||
651 sAspGlyLysAlaThrIleAsnLeuLysHisGlyGluSerLeuThrLeug 668
      ||| ||| :::::
7920 AGATGGCTCTAAATTCGTAAACACCTTACAGGTGCACCT..... 7959
      ::::: |||
668 InGlyLeuProGluGlyTyrSerTyrrLeuValLysGluThrAspSerGlu 684
      ::::: |||
7960 ..CAAGTAATAAAGTGGACGCTTCGTACTAAAGAGCTGTGCACAC 8007
      ::::: |||
685 .....GlyTyrrLysValLysValAsnSerGlnGluAlaAs 697
      ::::: |||
8008 GACAACATTAATTAATGAGATGCGCATTAACGAAACCAA..... 8049
      ::::: |||
697 nAlaThrValSerLysThrGlyIleThrSerAspGluThrLeuAlaPheG 714
      ::::: |||
8050 ..TCAACCATTCAGATGCCGTTGTAACAGATGACCCAAAC.....G 8089
      ::::: |||
714 LuAsn.....AsnLysGluProValIleProThrGlyVal 725
      ||| ||| :::::
8090 AAATACAGCTAATGTGGAAGATTCATTCACCTTATATCTACACTGT 8139
      ::::: |||
726 Asp 726
      |||
8140 GAT 8142
      ::::: |||
seq_name: /cgn1_7/ptodata/1/pna/us6006.COMB.seq:us-60-068-139-368
seq_documentation_block:
? Sequence 368, Application us/60068139
? GENERAL INFORMATION:
? APPLICANT: LAGACE, ROBERT E.
? APPLICANT: CORLEY, NEIL C.
? APPLICANT: RUSSO, FRANK D.
? APPLICANT: HANN, AMY L.
? APPLICANT: HEATH, JOE D.
? APPLICANT: FINNEY, GREGORY L.
? TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS
? TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USBS THEREOF
? NUMBER OF SEQUENCES: 471
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
? STREET: 3174 PORTER DRIVE
? CITY: PALO ALTO
? STATE: CALIFORNIA
? COUNTRY: USA
? ZIP: 94304
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk

```

```

/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/60/068,139
/ FILING DATE: HERewith
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: CERRONE, MICHAEL C.
/ REGISTRATION NUMBER: 39,132
/ REFERENCE/DOCKET NUMBER: PM-0008-2 P
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ. ID NO: 368:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 14346 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: circular
/ MOLECULE TYPE: genomic DNA
/ IMMEDIATE SOURCE:
/ CLONE: MCA1C368
/ US-60-068-139-368

```

```

alignment_scores:
  Quality: 169.50      Length: 864
  Ratio: 0.479        Gaps: 42
  Percent Similarity: 40.972  Percent Identity: 19.444

```

```

alignment_block:
  US-09-494-297-2 x US-60-068-139-368 ..

```

```

Align seq 1/1 to: US-60-068-139-368 from: 1 to: 14346

```

```

7 ProAsnLysLeuAsnThrLeuAsnThrGlnArgValLeuSerLysAsnSe 23
  |||:||||| |||:|||||
558 CCATCAGCAGTGATGACATTCAC.....CACCTTATTTTGATAAAATTC 601
23 rLysArgPheThrVal.....ThrLeuValGlyValPheLeuMet 37
  |||:||||| |||:|||||
602 CCCAAATATCCGATCTGCACCTGAGAAAACAGCGACGATGTGTTCATGC 651
37 lePheAlaLeuValThrSerMetValGlyAlaLysThrValPheGlyLeu 53
  |||:||||| |||:|||||
652 TAAATAAGCAATATACATCC.....AAATATATGTTAT 686
54 ValGluSerSerThrProAsnAlaIleAsnProAspSerSerSerGluTy 70
  |||:||||| |||:|||||
687 GGTGCATTTGATCACCCT..GCCAAAACCCA.....ACCTA 721
70 rArgTrpTyrGlyTyrGluSerTyrValArg.....GlyHisP 83
  |||:||||| |||:|||||
722 CATGAATTATCACAGAACAAAACATCAAAAACAAAACACGACGATG 771
83 rGlyArg.....TyrLysGlnPheArgValAlaHisAsp 93
  |||:||||| |||:|||||
772 ATTATCAAAACATTCTTTGGCTATATGGAGCTAAGA.....GAG 812
94 LeuArgValAsnLeuGluGlySerArgSerTyrGlnValTyrCysPheAs 110
  |||:||||| |||:|||||
813 CTGAGACCTAATAATAAAAAGTGCA..... 836
110 nLeuLysLysAlaPheProLeuGlySerAspSerSerValLysLysTrp 127
836 ..... 836
127 rLysLysHisAspGlyLysSerThrLysPheGluAspTyrAlaMetSer 143
836 ..... 836
144 ProArgIleThrGlyAspGluLeuAsnGlnLysLeuArgAlaValMetTy 160

```

```

837 ..... ||| :||| |||:|||||
160 rAsnGly.....HisProGlnAsnAlaAsnGlyIleMetG 172
  |||:||||| |||:|||||
872 CACCACACCTACTTATTTATTCATGCTGAGAAATGCCAGCCACTCTGC 921
172 LucGlyLeuGluProLeuAsnAlaIleArgValThrGlnGluAlaValTyr 188
  |||:||||| |||:|||||
922 CA.....AAGCGCGTAAATTGTGACTGTAGGGCAATGG 956
189 TyrTyrSerAspAsnAlaPro.....IleSerAsnProAspG 201
  |||:||||| |||:|||||
957 TTGTATCTGACCATGTCGCAAAAACGCCCATTTTGTAGTAAACAGACGA 1006
201 u.....SerPheLysArgGluSerGluSerAsnL 211
  |||:||||| |||:|||||
1007 TAAAGTAGGCACCTTATTTTATCTACACCAAGAAATCAAAATGAAGGCGATT 1056
211 euValSerThrSerGlnLeuSerLeu..... 219
1057 TGGTGAATGACGACGACATTTATCTAACAGCTTAATATAACACACC 1106
220 .....MetArgGlnAlaLe 224
1107 CCGGCACATTATAGCGTGACCTTTGATCAAAATACCTTAAAGCAATAT 1156
224 uLysGlnLeuIleAspProAsnLeuAlaThrLysMetProLysGlnValP 241
  |||:||||| |||:|||||
1157 GCTCTTATTATGACACCAACCAACAAACAGACCGGATGGCTTATATATCA 1206
241 roAspAspPhe.....GlnLeuSerIlePheGlu 250
  |||:||||| |||:|||||
1207 GAAGTCAGTTTGTATACCGACAAAAGGTCAATGAGCGCATGTCTATGAG 1256
251 SerGluAspLys.....GlyAspLysTyrAsnGlyTyrGlnAsnLe 265
  |||:||||| |||:|||||
1257 ATTGACGCCAAGATTATGTCACACCGCTTACTGCGCACACCAAACTCT 1306
265 uLeuSerGlyGlyLeu.....ValProThrL 274
  |||:||||| |||:|||||
1307 GATTGATGATTAACACCAATACCGACCTTTTGTAAAGACTGTCTCCA 1356
274 yspProProThrProGlyAspProPheMetProProAsnGlnProGlnThr 290
  |||:||||| |||:|||||
1357 AAAAAGCCATCCCAACACCA.....GACCCCAAC..... 1388
291 ThrSerValLeuIleArgLysTyrAlaIleGlyAspTyrSerLysLeuLe 307
  |||:||||| |||:|||||
1389 .....TCAGATACGCT 1399
307 uGluGlyAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheGlnA 324
  |||:||||| |||:|||||
1400 AGAAGCGCGGCTTTTANGTGAGTGGGCGATGAGCTG.....G 1437
324 laArgValPheSerSerAsnAspIle.....Gly 333
  |||:||||| |||:|||||
1438 CGGTAATTTTATTCATGACAAACCACTTTTGTGCTTGTGGTGGC 1487
334 GluArgIleGluLeuSerAspGlyThrTyrThrLeuThrGluLeuAsnSe 350
  |||:||||| |||:|||||
1488 AAACGAGACAAAACGACGAAACCTGTCGCCCAAAAACGGGTGATTTTGA 1537
350 rProAlaGlyTyrSerIleAlaGluProIleThrPheLysValGluAlaG 367
  |||:||||| |||:|||||
1538 T...ACAGGATTGAA.....AAACCCAGACACAGCTTTGTTGGCCATG 1578
367 LysValTyrThrIleIleAspGlyLys.....GlnIleGlu 379
  |||:||||| |||:|||||
1579 AAGGATTTGGATGATTTATGACGCTAAAGTTAATGATGAAGTCAAT 1628
380 AsnProAsnLysGluIleValGluProTyrSerValGluAlaTyrAsnAs 396
  |||:||||| |||:|||||

```

1629	AATGCAATTTGAAGATGAACAACTGTCCTCCTGCAATATAAGAAATATATATGA	1678
396	ppe.....GluIupheserValI	403
1679	ATATATATTTATGAGACACCAACAAATTCACCAAAAAATAAAGCGCA	1728
403	euthrThgIlnasntyrAla.....LysPheTyrTyr	413
1729	GGGTCCAAAAAACCCCTGCTATTTTGGTCAGCATGATTAAGCTTTATATTTT	1778
414AlaLys.....AsnLysasnGlyIse	420
1779	AATGTAACATATTATGACTTATCACCACAAAGACAAACAAAGCTTGGCT	1828
420	rSerGlnValValTyrCysPheasnAlaAspLeuLysSerProProAsps	437
1829	CTCCCAAGATACACGACACCAATAAAGATATTTTGGCTAAATATCCOAGATG	1878
437	ergIuAspGlyGlyLysThrMetThrProAspPheThr.....	450
1879	CC.....AAAGTAAGCAGACAAATAAAGTTACCAAAATGCTT	1916
450	450
1917	CTACAAACAAGCCAAAGATTAAGCCGATATACCGCATTCATGCCAAAAGCTA	1966
451GlyIuValLysTyrThrHisIleAlaGlyArg	462
1967	TGACCAACATCACTTTTGGGAAGTATTTGTAATTAATATACAAAGCAAC	2016
462	spleuPhelystyrThrValLysProArgAspThrAspProAspThrPhe	478
2017	CACACGCGAGTTATTTTGGCAAGCGGCTCAAGCGATGTACATACCTCAG	2066
479	IeuLys.....HisIiI	482
2067	GTGCCCACTGACGGTAATTCACCTATTAATAGCTTTGGCAGCGCTACTT	2116
482	eLysLysValIleGlyLysGlyTyrArgGluLysGlyIlnaIleGluT	499
2117	GACCCGAAAAAAGCAAAAGTTTATGCAAAAGTACAGATTCACATCAACG	2166
499	yrSerGlyLeuThrGluThrGluLeuArgAlaIleThrGluLeuAlaIle	515
2167	AAAAAGCTTATAAAT.....	2183
516	TyrTyrPheThr.....AspSerAlaGluLeuAs	525
2184	TATATATTTACCAAAAGACTTATTCACACAGATGACGATGACGATGACGA	2233
525	plyAspLysIleu.....LysAspTyrHisG	534
2234	TGACGATACATTTGACCGCATCTGATGATTCACACAGATATTAATACACTG	2283
534	LysPheGlyAspMetLsnAspSerThrIleuAlaValAlaLysIleLeuVal	550
2284	GGGATGATGATTTG.....ATTGCATCT	2306
551	GluTyrAlaGlnAspSerAsnProProGluLeuThrAspLeuAspPheP	567
2307	GATGATTTCAAGATGATGACCCAGATGGCAGTACGATTCAGATGATTT	2356
567	eIleProAnaAnaAsnLysTyrGlnSerLeuIleGlyThrGlnTrpHisP	584
2357	G.....GGTGATGGTGCAGATGATGACCGCGCAGCAAAAGTATATCTG	2400
584	rogluAspLeuValAspIleIleArgMetGluAspLysGlyGluValIle	600
2401	CAGGTAAAT.....ATTGCCCTGGAATTTGAAAACAATAATACCTG	2438
601	ProVal.....ThrHisAsnLeuThrIleu.....	608
2439	CCCATTAATTAAGACCTATCTCATAAAAAAGCTTTGGCCCTAGATGTAATAAA	2488

[illegible]

```

117 .....LeuGlySerAspSerSerValLysLysTPTPTLysLys 129
1757 ATATATCTTCAATTACGAGCTACAAAGAAATTTATTTTATTACAAAG... 1803
130 HisAspGlyIleSerThrLysPheGluAspTyrLamSerProArg11 146
1804 .....TTAGSAGC.....GATTATACAGTAACGCCAAGCTC 1835
146 ethrGlyAspGluLeuAsnGlnLysLeuArgAlaValMetLysAsnGly 163
1836 AGATGGTTCAGTTATTAAGTTCACCTACGCCAATACCAACGAAATCCAAA 1885
163 .IsProGlnAsnAlaAsnGlyIleMetGluGluLeuGlu..... 175
1886 TTCCAATTGGTTTAAATTATGTCGCAAGATAGTTGTCACAAAAGATAAAAGT 1935
176 ..ProLeuAsnAlaIleArgValThrGlnGluAlaValTPTPTyrSe 191
1936 ATCCCACTTCGATACGATACGATACAAATGAGTCTGAGSTTTAACTCC 1985
191 rAspAsnAlaProIleSerAsnProAspGluSerPheLysArgLysSerg 208
1986 AGTTGATACACAGTAAGTACTACT.....AATAGTAAGCGGTGTTCTG 2026
208 lUserAsnLeuValSerThrSergLneSerLneMetArgGlnAlaLeu 224
2027 AACGACACCTTCAAGTAGTAAATCAATTCCTTGCATACGACGAAT 2076
225 LysGlnLeuIleAspProAsnLeuAlaThrLysMetProLysGlnValPr 241
2077 GATCTTTTGATCCTACATAAGGCGTCGTACAAATTCACAGCTGCGCC.. 2124
241 oAspAspPheGlnLeuSerIlePheGluSergLysAspLysAspLysT 258
2125 ....GATGTTCTTTTGCATTTATGATGTTTCAAC.....GATCAGG 2164
258 YrAsnLysGlyTyrGlnAsnLeuLeuSergLysGluLeuValProThrLys 274
2165 TAGATTTCAATTTATCCAAATACACTGGGACCGCGGTCAATCTTGATGAAA 2214
275 ProProThrProGlyAspProProMetPro.....ProAsnGlnProG1 289
2215 CCAATGACGCCAAACAGCCCTGGATATCCAAAGATTTACTTTGACGAAA 2264
289 nThrThrServalLeuIle.....ArgLysTyrAlaI 300
2265 TACCAATAGTTACACGTTTGTGAAAAAACCAACAAACGTTACATTA 2314
300 lGlyAspTyrSerLysLeu.....LeuGluGlyAlaThrLeuGln 313
2315 TT...GAGTTAAAAACGCCAATGCTGATCGACGTCGCCAATCTTTAT 2361
314 LeuThrGly.....AspAsnValAsnSerPheGlnAl 324
2362 ATACACAGGACAGCAAGAAACACCAATCGAATATATATGAAGCTCTGC 2411
324 aArgValPheSerSerAsnAspIleGlyGluArgIleGluLeuSerSpg 341
2412 TTCGTTTCTGTTCAAAAT.....GAAGCTTTAGACATTTTGAAGTG 2452
341 lYrThrTyrThrLeuThrGluLeuAsnSerProAlaGlyTyrSerIleAla 357
2453 CAACA.....CAAGCGCG 2466
358 GluProIleThrPheLysValGluAlaGlyLysValTyrThr...Ile11 373
2467 AATCCAAACATTAAAAAATGTAAACAAAACGACATTAACAAAAAATAT 2516
373 eAspGlyLys.....GlnIleGluAsnProAsnLysGluIleValGluP 388
2517 TGATTAATAAAACACATCGTGTGAATAATCCAAAGATTAATTAACACAA 2566

```

```

388 roTyrServalGluAlaTyrAsnAspPheGluGluPheServalLeuThr 404
2567 AAGCACAACCAATGCTCAATATGATTTGAATTCATACCGTG..... 2610
405 ThrGlnAsnTyrAlaLysPheTyrTyrAlaLysAsnLysAsnGlySerSe 421
2611 .....AAAGCGTGC 2621
421 rGlnValValTyrCysPheAsnAlaAspLeuLysSerProProAspSerg 438
2622 AGAAGATGCTTAT.....TCATTGACAGACACTA 2650
438 lAspArgGlyLysThrMetThrProAspPheThrThrGlu..... 452
2651 CAACGCGTCCGAAGTTCATTTTAAAGACTATACATTGACAGAAACATTT 2700
453 ..ValLysTyrThrHisIleAlaGlyArgAspLeuDeuThelystyrThva 468
2701 ACGATTGAATACATTAACGCTGTCGAAACGCTGCCAAATCTTATACAGA 2750
468 lLysProArgAspThrAspProAspThrPheLeuLysHisIleLysLysY 485
2751 AACAAACATGACCTGTGAACATTTGAACAGATGCTGTAGCAAGAAA 2800
485 all.....GluLys 488
2801 AAGTCACACACTGCCCAATGACATTTGAATTCGAGAAGTGATGCCGA 2850
489 GlyTyrArgGluLysGlyGlnAlaIleGluTyrSergLysLeuThrGluTh 505
2851 GGATTTGTTTATTATGCAACTGCCACATTTCTACACGATTAACGAGAGA 2900
505 rGlnLeuArgAlaIleThrGlnLeuAlaIleTyrTyrPheThrAsp... 520
2901 TGAACCAAGCAATTCGAAAGTTCTTTGAACATTAATGATATGTCA 2950
521 .....SerAlaGluLeuAspLysAspLysLeuLysAspTyrHisGly 534
2951 CGCATACAGCACCGCAATTTACACAGATGAGAAAAGTCAATCTCTTT 3000
535 PheGlyAspMet..AsnAspSerThrLeuAlaValAlaLysIleLeuVa 550
3001 GATGCCATCATGACACGCTGATTACTTTGCGAGTAACGATTAACGCCGA 3050
550 lGluTyrAlaGlnAspSerAsnProGlnLeuThr..... 562
3051 GGAATATTCGTTGATGAAGAG.....TATTTGACAGAAAGCCATT 3094
563 .....AspLeuAspPhePheIlePro..... 569
3095 AGCTGTCMAAGAGACACCACTAAATTCATTAACGAAACAAAT 3144
570 AsnAsnAsnLysTyrGln.....SerLeuIleGlyThrG1 581
3145 GATCACAGCTGTTTACAAAGTCAAGATTCAAAGATTTATGTGCGCGATT 3194
581 nThrHisProGluAspLeuValAspIleIleArgMetGluAspLysLys. 597
3195 ATGGAAACCAAGAGAG.....AAGTTGTTTACAGCAACGATTAACAG 3238
598 ..GluValIleProValThrHisAsnLeuThrLeuArgLysThrValThr 613
3239 GTCAAAGACGTTCCCTTCGA.....AAATACAGGTTTCA 3273
614 GlyLeuAlaGlyAspArgThrLysAspPheHisPheGluIleGluLeuLys 630
3274 GGTCAAGTT...GATAACACCAAGAGCGGCTTATCCAAATTTATTTACAG 3320
630 sAsn..AsnLysGlnGluLeuLeuSergLynThrValLysThrAspLysT 646
3321 TGAGGAAGGTAAAGAAAGAACAGCTATGTGACCGTGAACCCGCAAT 3370
646 hYrAsnLeuGluPheLysAspGlyLysAlaThrIleAsnLeuLysHisGly 662

```



```

3371 CTAAGTAGAGGTCAAAGAT.....ACAAGATTATGTT.....GGT 3408
663 GluSerLeuThrLeuGlnGlyLeuProGluGlyTyrSerTyrLeuValIy 679
3409 GATTCGTGCGAAA.....CCAGAA.....GATAATTTCGTTTC 3440
679 sgluthraspSergIuGlyTyrIysVal.....LysValAsnSerg 693
3441 AGCGACAGACAAACAGGTCAAGACGTTCCGTTGAAAAAATTGATGTTG 3490
693 InGluValAlaAsnAlaThrValSerLysThrGlyTleThrSerAspGlu 709
3491 AGGGAACAGTCAAT.....GTTGATAAATAGGC.....GATTAT 3525
710 ThrLeuAlaPheGluAsnAsnLysGlu 718
3526 GAAATTGTCTATAAAATGCCAAAAA 3552

```

